

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representation of
The original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:54 : Search time 9.81254 Seconds
(without alignments)
4281.862 Million cell updates/sec

Title: US-09-835-992A-19

Perfect score: 218
Sequence: 1 cgcagagaaagttatttca.....cacagtgttgcacattca 714

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_n2p.model -DEV=olp
-Q=/cgn2_1/USPTO_spool/US09835992/runat_14012003_161527_4050/app_query.fasta_1.3932
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=olig.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992@cgn_1_1_49@runat_14012003_161527_4050 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	3.2	18	6 5182195-2	Patent No. 5182195
2	7	3.2	19	4 US-09-149-476-534	Sequence 534, App
3	7	3.2	21	6 5182195-7	Patent No. 5182195
4	7	3.2	44	4 US-09-288-143-88	Sequence 88, Appl
5	7	3.2	62	1 US-08-488-961-6	Sequence 6, Appl1
6	7	3.2	62	4 US-08-973-297-6	Sequence 6, Appl1
7	7	3.2	62	5 PCT-US96-06511-6	Sequence 6, Appl1
8	7	3.2	89	4 US-09-149-476-451	Sequence 451, App
9	7	3.2	187	4 US-09-149-476-417	Sequence 417, App
10	7	3.2	222	1 US-07-869-933-12	Sequence 12, Appl
11	7	3.2	222	1 US-07-869-933-28	Sequence 28, Appl
12	7	3.2	222	4 US-09-103-663-12	Sequence 12, Appl

13	7	3.2	222	4	US-09-103-663-28	Sequence 28, Appl
14	7	3.2	267	2	US-08-557-128-4	Sequence 4, Appl1
15	7	3.2	267	4	US-09-242-690A-36	Sequence 36, Appl1
16	7	3.2	326	1	US-07-603-133B-25	Sequence 25, Appl1
17	7	3.2	380	4	US-08-857-076-110	Sequence 110, App
18	7	3.2	420	5	PCT-US96-10602-10	Sequence 10, Appl
19	7	3.2	427	4	US-09-134-001C-5143	Sequence 5143, Ap
20	7	3.2	554	4	US-09-321-276-4	Sequence 4, Appl1
21	7	3.2	554	4	US-08-916-481-2	Sequence 2, Appl1
22	7	3.2	563	4	US-08-916-481-3	Sequence 3, Appl1
23	7	3.2	663	4	US-09-196-293-5	Sequence 5, Appl1
24	7	3.2	663	4	US-08-209-603E-5	Sequence 5, Appl1
25	7	3.2	663	4	US-08-235-836C-70	Sequence 70, Appl
26	7	3.2	663	4	US-08-235-836C-78	Sequence 78, Appl
27	7	3.2	681	4	US-08-760-615-4	Sequence 4, Appl1
28	7	3.2	681	4	US-08-760-615-6	Sequence 6, Appl1
29	7	3.2	700	4	US-08-235-836C-74	Sequence 74, Appl
30	7	3.2	1698	4	US-09-315-793-12	Sequence 12, Appl
31	7	3.2	1724	4	US-08-857-076-12	Sequence 12, Appl
32	6	2.8	11	1	US-08-338-634-4	Sequence 4, Appl1
33	6	2.8	11	1	US-08-787-547-67	Sequence 67, Appl
34	6	2.8	11	5	PCT-US95-16415-1	Sequence 1, Appl1
35	6	2.8	14	1	US-08-232-453A-40	Sequence 40, Appl
36	6	2.8	15	4	US-09-027-900-3	Sequence 3, Appl1
37	6	2.8	17	2	US-08-497-599-2	Sequence 2, Appl1
38	6	2.8	17	2	US-08-497-599-18	Sequence 18, Appl
39	6	2.8	18	4	US-08-847-844A-50	Sequence 50, Appl
40	6	2.8	20	4	US-09-513-783A-146	Sequence 146, App
41	6	2.8	21	3	US-08-865-297-8	Sequence 8, Appl1
42	6	2.8	21	3	US-08-865-297-11	Sequence 11, Appl
43	6	2.8	21	3	US-08-946-329A-40	Sequence 40, Appl
44	6	2.8	22	1	US-08-241-853-37	Sequence 37, Appl
45	6	2.8	22	1	US-08-241-853-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
Patent No. 5182195
APPLICANT: NAKAHAMA, KAZUO; KAISHO, YOSHIHIKO; YOSHIMURA, KOJI
TITLE OF INVENTION: METHOD FOR INCREASING USING PROTEASE
DEFICIENT YEASTS
NUMBER OF SEQUENCES: 71
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/269,140
FILING DATE: 09-NOV-1988
SEQ ID NO: 2:
LENGTH: 18
5182195-2

Alignment Scores:

Pred. No.: 123
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.21%
DB: 6
Gaps: 0

US-09-835-992A-19 (1-714) x 5182195-2 (1-18)

QY 493 CTGTTGGCATTTGTTCTTA 513

Db 6 LeuLeuAlaLeuCysPheLeu 12

RESULT 2

US-09-149-476-534
Sequence 534, Application US/09149476

Patent No. 6420526

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002P1

1	CURRENT	APPLICATION	NUMBER:	US/09/149,476
2	CURRENT	FILING DATE:	1998-09-08	
3	EARLIER	APPLICATION	NUMBER:	PCT/US98/04493
4	EARLIER	APPLICATION	NUMBER:	1998-03-06
5	EARLIER	APPLICATION	NUMBER:	60/040,162
6	EARLIER	FILING DATE:	1997-03-07	
7	EARLIER	APPLICATION	NUMBER:	60/040,333
8	EARLIER	APPLICATION	NUMBER:	1997-03-07
9	EARLIER	APPLICATION	NUMBER:	60/038,621
10	EARLIER	FILING DATE:	1997-03-07	
11	EARLIER	APPLICATION	NUMBER:	60/040,626
12	EARLIER	FILING DATE:	1997-03-07	
13	EARLIER	APPLICATION	NUMBER:	60/040,334
14	EARLIER	FILING DATE:	1997-03-07	
15	EARLIER	APPLICATION	NUMBER:	60/040,336
16	EARLIER	FILING DATE:	1997-03-07	
17	EARLIER	APPLICATION	NUMBER:	60/040,163
18	EARLIER	FILING DATE:	1997-03-07	
19	EARLIER	APPLICATION	NUMBER:	60/047,600
20	EARLIER	FILING DATE:	1997-05-23	
21	EARLIER	APPLICATION	NUMBER:	60/047,615
22	EARLIER	FILING DATE:	1997-05-23	
23	EARLIER	APPLICATION	NUMBER:	60/047,597
24	EARLIER	FILING DATE:	1997-05-23	
25	EARLIER	APPLICATION	NUMBER:	60/047,502
26	EARLIER	FILING DATE:	1997-05-23	
27	EARLIER	APPLICATION	NUMBER:	60/047,633
28	EARLIER	FILING DATE:	1997-05-23	
29	EARLIER	APPLICATION	NUMBER:	60/047,583
30	EARLIER	FILING DATE:	1997-05-23	
31	EARLIER	APPLICATION	NUMBER:	60/047,617
32	EARLIER	FILING DATE:	1997-05-23	
33	EARLIER	APPLICATION	NUMBER:	60/047,618
34	EARLIER	FILING DATE:	1997-05-23	
35	EARLIER	APPLICATION	NUMBER:	60/047,503
36	EARLIER	FILING DATE:	1997-05-23	
37	EARLIER	APPLICATION	NUMBER:	60/047,592
38	EARLIER	FILING DATE:	1997-05-23	
39	EARLIER	APPLICATION	NUMBER:	60/047,581
40	EARLIER	FILING DATE:	1997-05-23	
41	EARLIER	APPLICATION	NUMBER:	60/047,584
42	EARLIER	FILING DATE:	1997-05-23	
43	EARLIER	APPLICATION	NUMBER:	60/047,500
44	EARLIER	FILING DATE:	1997-05-23	
45	EARLIER	APPLICATION	NUMBER:	60/047,587
46	EARLIER	FILING DATE:	1997-05-23	
47	EARLIER	APPLICATION	NUMBER:	60/047,492
48	EARLIER	FILING DATE:	1997-05-23	
49	EARLIER	APPLICATION	NUMBER:	60/047,598
50	EARLIER	FILING DATE:	1997-05-23	
51	EARLIER	APPLICATION	NUMBER:	60/047,613
52	EARLIER	FILING DATE:	1997-05-23	
53	EARLIER	APPLICATION	NUMBER:	60/047,582
54	EARLIER	FILING DATE:	1997-05-23	
55	EARLIER	APPLICATION	NUMBER:	60/047,596
56	EARLIER	FILING DATE:	1997-05-23	
57	EARLIER	APPLICATION	NUMBER:	60/047,612
58	EARLIER	FILING DATE:	1997-05-23	
59	EARLIER	APPLICATION	NUMBER:	60/047,632
60	EARLIER	FILING DATE:	1997-05-23	
61	EARLIER	APPLICATION	NUMBER:	60/047,601
62	EARLIER	FILING DATE:	1997-05-23	
63	EARLIER	APPLICATION	NUMBER:	60/043,580
64	EARLIER	FILING DATE:	1997-04-11	
65	EARLIER	APPLICATION	NUMBER:	60/043,568
66	EARLIER	FILING DATE:	1997-04-11	
67	EARLIER	APPLICATION	NUMBER:	60/043,314
68	EARLIER	FILING DATE:	1997-04-11	
69	EARLIER	APPLICATION	NUMBER:	60/043,569
70	EARLIER	FILING DATE:	1997-04-11	
71	EARLIER	APPLICATION	NUMBER:	60/043,311
72	EARLIER	FILING DATE:	1997-04-11	
73	EARLIER	APPLICATION	NUMBER:	60/043,671

EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/043,674
EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/043,669
EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/043,312
EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/043,313
EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/043,672
EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/043,315
EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/048,974
EARLIER FILING DATE:	1997-06-06
EARLIER APPLICATION NUMBER:	60/056,886
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,877
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,889
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,893
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,630
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,878
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,662
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,872
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,882
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,637
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,903
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,888
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,879
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,880
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,894
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,911
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,636
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,874
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,910
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,864
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,631
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,845
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056,892
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/057,761
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/047,595
EARLIER FILING DATE:	1997-05-23
EARLIER APPLICATION NUMBER:	60/047,599
EARLIER FILING DATE:	1997-05-23
EARLIER APPLICATION NUMBER:	60/047,588
EARLIER FILING DATE:	1997-05-23
EARLIER APPLICATION NUMBER:	60/047,585
EARLIER FILING DATE:	1997-05-23
EARLIER APPLICATION NUMBER:	60/047,586
EARLIER FILING DATE:	1997-05-23

EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Alignment Scores:

Pred. No.:	122	Length:	19
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.21%	Indels:	0
DB:	4	Gaps:	0

US-09-835-992A-19 (1-714) x US-09-149-476-534 (1-19)

OY 624 TTACTGTCCTAATTATTC 644
Db 4 LeuLeuPheLeuLeuLeu 10

RESULT 3

Patent No. 5182195
APPLICANT: NAKAHAMA, KAZUO;KAISHO, YOSHIHIKO;YOSHIMURA, KOJI
TITLE OF INVENTION: METHOD FOR INCREASING USING PROTEASE
DEFICIENT YEASTS
NUMBER OF SEQUENCES: 71
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/269,140
FILING DATE: 09-NOV-1988

SEQ ID NO:7:
LENGTH: 21
5182195-7

Alignment Scores:

Pred. No.:	120	Length:	21
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.21%	Indels:	0
DB:	6	Gaps:	0

US-09-835-992A-19 (1-714) x 5182195-7 (1-21)

OY 493 CTGTGGCATTGTTCTTA 513
Db 6 LeuLeuAlaLeuCysPheLeu 12

RESULT 4

US-09-288-143-88
Sequence 88, Application US/09288143
Patent No. 6433139

GENERAL INFORMATION:

APPLICANT: Brewer et al.
TITLE OF INVENTION: 53 Human Secreted Proteins
FILE REFERENCE: P2018P1
CURRENT APPLICATION NUMBER: US/09/288,143
CURRENT FILING DATE: 1999-04-08
EARLIER APPLICATION NUMBER: PCT/US98/21142
EARLIER FILING DATE: 1998-10-08
EARLIER APPLICATION NUMBER: 60/061,463
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,529
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/071,498
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,527
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,536
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,532
EARLIER FILING DATE: 1997-10-09
NUMBER OF SEQ ID NOS: 219
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 88
LENGTH: 44
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (44)
OTHER INFORMATION: Xaa equals stop translation

US-09-288-143-88

Alignment Scores:

Pred. No.:	107	Length:	44
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.21%	Indels:	0
DB:	4	Gaps:	0

US-09-835-992A-19 (1-714) x US-09-288-143-88 (1-44)

OY 621 ATTTACTGTCCTAATTTA 641
Db 5 IleLeuPheLeuLeuLeu 11

RESULT 5

US-08-488-961-6
Sequence 6, Application US/08488961
Patent No. 5606042
GENERAL INFORMATION:

APPLICANT: Smith, Daniel S.
APPLICANT: Walker, John C.
TITLE OF INVENTION: Glycine and Phaseolus
TITLE OF INVENTION: alpha-D-Galactosidases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Relisting, Ethington, Barnard & Perry
STREET: P.O. Box 4390
CITY: Troy
STATE: Michigan
COUNTRY: US
ZIP: 48099-4390
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,961
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: P-320 (UMO)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 689-3500
TELEFAX: (810) 689-4071
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 62 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-488-961-6

Alignment Scores:
Pred. No.: 101
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.21%
DB: 1
Gaps: 0

US-09-835-992A-19 (1-714) x US-08-488-961-6 (1-62)

QY 496 TTGCATTGTCTTCTTACTT 516
Db 20 leuAlaLeucysPheLeuLeu 26

RESULT 6
US-08-973-297-6
Sequence 6, Application US/08973297
Patent No. 6184017
GENERAL INFORMATION:
APPLICANT: Smith, Daniel S.
APPLICANT: Walker, John C.
TITLE OF INVENTION: Glycine and Phaseolus
TITLE OF INVENTION: alpha-D-Galactosidases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 6184017thwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/973,297
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0994.00050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 539-5050
TELEFAX: (810) 539-5055
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 62 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-973-297-6

Alignment Scores:
Pred. No.: 101
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.21%
DB: 4
Gaps: 0

US-09-835-992A-19 (1-714) x US-08-973-297-6 (1-62)

QY 496 TTGCATTGTCTTCTTACTT 516
Db 20 leuAlaLeucysPheLeuLeu 26

RESULT 7
PCT-US96-06511-6
Sequence 6, Application PC/TUS9606511
GENERAL INFORMATION:
APPLICANT: Smith, Daniel S.
APPLICANT: Walker, John C.
TITLE OF INVENTION: Glycine and Phaseolus
TITLE OF INVENTION: alpha-D-Galactosidases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 Northwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06511
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0994.00050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 539-5050
TELEFAX: (810) 539-5055
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 62 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US96-06511-6

Alignment Scores:

Pred. No.: 101 Length: 62
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 5 Gaps: 0

US-09-835-992A-19 (1-714) x PCT-US96-06511-6 (1-62)

OY 496 TTGGCATTCGTCTTCTTACTT 516

Db 20 LeuAlaLeuCysPheLeuLeu 26

RESULT 8

US-09-149-476-451

; Sequence 451, Application US/09149476

; Patent No. 6420526

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 186 Human Secreted proteins

; FILE REFERENCE: P2002P1

; CURRENT APPLICATION NUMBER: US/09/149,476

; CURRENT FILING DATE: 1998-09-08

; EARLIER APPLICATION NUMBER: PCT/US98/04493

; EARLIER FILING DATE: 1998-03-06

; EARLIER APPLICATION NUMBER: 60/040,162

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,333

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/038,621

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,626

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,334

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,336

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,163

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/047,600

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,615

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,597

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,502

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,633

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,583

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,617

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,618

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,503

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,592

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,581

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,584

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,500

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,587

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,492

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,598

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,613

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/056,910

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/047,582

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,596

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,612

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,632

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,601

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/043,580

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,568

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,314

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,569

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,311

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,671

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,674

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,669

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,312

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,313

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,672

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,315

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/048,974

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/056,886

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,877

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,889

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,893

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,630

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,878

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,662

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,872

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,882

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,637

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,903

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,888

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,879

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,880

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,894

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,911

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,636

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,874

; EARLIER FILING DATE: 1997-08-22

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Alignment Scores: 95.7 Length: 89
Pred. No.: 7.00 Matches: 7
Score:

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 4 Gaps: 0
US-09-835-992A-19 (1-714) x US-09-149-476-451 (1-89)
QY 537 CTTAATTGAATCTTCAGCAG 557
Db 62 LeuLysLeuAsnLeuGlnGln 68
RESULT 9
US-09-149-476-417
Sequence 417, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
Applicant: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002p1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Alignment Scores: 85
Pred. No.: 7.00
Score: 100.00%
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Length: 187
Matches: 7
Conservative: 0
Mismatch: 0

Query Match: 3.23% Indels: 0
DB: 4 Gaps: 0
US-09-835-992A-19 (1-714) x US-09-149-476-417 (1-187)
QY 164 TTGTTAATTTTCCTTTT 144
DB 6 PheValIleIlePheIleuphe 12
RESULT 10
US-07-869-933-12
; Sequence 12, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Rat
; STRAIN: FCRI alpha subunit
; US-07-869-933-12
Alignment Scores:
Pred. No.: 82.8 Length: 222
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
Gaps: 1
US-09-835-992A-19 (1-714) x US-07-869-933-12 (1-222)
QY 83 AATAACATTAGTATTAGAAA 103
DB 134 AsnAsnIleSerIleArgLys 140
RESULT 11
US-07-869-933-28
; Sequence 28, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:

APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; STRAIN: alpha subunit
; US-07-869-933-28
Alignment Scores:
Pred. No.: 82.8 Length: 222
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
Gaps: 1
US-09-835-992A-19 (1-714) x US-07-869-933-28 (1-222)
QY 83 AATAACATTAGTATTAGAAA 103
DB 134 AsnAsnIleSerIleArgLys 140
RESULT 12
US-09-103-663-12
; Sequence 12, Application US/09103663D
; Patent No. 6171803
; GENERAL INFORMATION:
; APPLICANT: Kinet et al.
; TITLE OF INVENTION: Isolation, characterization, and use of the human beta
; TITLE OF INVENTION: subunit of the high affinity receptor for
; TITLE OF INVENTION: immunoglobulin E.
; FILE REFERENCE: 50490
; CURRENT APPLICATION NUMBER: US/09/103,663D
; CURRENT FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 07/869,933
; EARLIER FILING DATE: 1992-04-16
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Rattus sp.

US-09-103-663-12

Alignment Scores:
Pred. No.: 82.8 Length: 222
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-103-663-12 (1-222)

QY 83 AATAACATTACTATTAGAAAA 103

Db 134 AsnAsnIleSerIleArgLys 140

RESULT 13

US-09-103-663-28

; Sequence 28, Application US/09103663D

; Patent No. 6171803

; GENERAL INFORMATION:

; APPLICANT: Kinet et al.

; TITLE OF INVENTION: Isolation, characterization, and use of the human beta

; TITLE OF INVENTION: subunit of the high affinity receptor for

; TITLE OF INVENTION: Immunoglobulin E.

; FILE REFERENCE: 50490

; CURRENT APPLICATION NUMBER: US/09/103,663D

; CURRENT FILING DATE: 1998-06-23

; EARLIER APPLICATION NUMBER: 07/869,933

; EARLIER FILING DATE: 1992-04-16

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 28

; LENGTH: 222

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-103-663-28

Alignment Scores:
Pred. No.: 82.8 Length: 222
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-103-663-28 (1-222)

QY 83 AATAACATTACTATTAGAAAA 103

Db 134 AsnAsnIleSerIleArgLys 140

RESULT 14

US-08-557-128-4

; Sequence 4, Application US/08557128

; Patent No. 5849524

; GENERAL INFORMATION:

; APPLICANT: KONDO, Keiji

; APPLICANT: KAJIWARA, Susumu

; APPLICANT: MISAWA, No. 5849524hiko

; TITLE OF INVENTION: TRANSFORMATION SYSTEMS FOR THE YEAST

; TITLE OF INVENTION: CANDIDA UTILIS AND THE EXPRESSION OF HETEROLOGOUS GENES

; TITLE OF INVENTION: THEREWITH

; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/557,128

FILING DATE: 25-JAN-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/JP95/01005

FILING DATE: 25-MAY-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-129287

FILING DATE: 28-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-285823

FILING DATE: 26-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-135015

FILING DATE: 25-MAY-1994

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 49441/108

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-557-128-4

Alignment Scores:
Pred. No.: 80.4 Length: 267
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-19 (1-714) x US-08-557-128-4 (1-267)

QY 521 GACTAGTACAGAACACAAT 501

Db 80 GluLeuSerLysIshisAsn 86

RESULT 15

US-09-242-690A-36

; Sequence 36, Application US/09242690A

; Patent No. 6284534

; GENERAL INFORMATION:

; APPLICANT: KONDO, KEIJI

; APPLICANT: MIURA, YUTAKA

; TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE

; TITLE OF INVENTION: SAME

; FILE REFERENCE: 049441/0118

; CURRENT APPLICATION NUMBER: US/09/242,690A

; CURRENT FILING DATE: 1999-02-23

; PRIOR APPLICATION NUMBER: PCT/JP97/02924

; PRIOR FILING DATE: 1997-08-22

; PRIOR APPLICATION NUMBER: JP 8/241062

; PRIOR FILING DATE: 1996-08-23

; NUMBER OF SEQ ID NOS: 66

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 36

; LENGTH: 267

; TYPE: PRT

; ORGANISM: Candida utilis

US-09-242-690A-36

Alignment Scores:

Pred. No.:	80.4	Length:	267
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.23%	Indels:	0
DB:	4	Gaps:	0

US-09-835-992A-19 (1-714) x US-09-242-690A-36 (1-267)

QY 521 GAACTAAGTAAGAAACACAAAT 501
|||||
Db 80 GluLeuSerLysLysHisAsn 86

Search completed: January 14, 2003, 17:42:08
Job time : 12.0625 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:24 ; Search time 35.1708 Seconds
(without alignments)
5410.226 Million cell updates/sec

Title: US-09-835-992A-19
Perfect score: 218
Sequence: 1 cgcacgagaaagttatttca.....cacagttgttgcaattca 714

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+ -n2p.model -DEV-xlp
-O=/cgn2_1/USPRO_spool/US09835992/runat_14012003_161525_3979/app_query.fasta_1.3932
-DB-A_Geneseq_101002 -QFMT-fastan -SUFFIX-olig.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=-1 -MATRIX-oligo -TRANS-human40.cdl
-LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE-LOCAL
-OUTFMT-pto -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER-US09835992_@CGN_1_1_121_@runat_14012003_161525_3979 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	8	3.7	29	AAW40006	Peptide effecting
2	8	3.7	159	ABB31093	Peptide #3744 enco
3	8	3.7	159	ABB36288	Peptide #3794 enco
4	8	3.7	159	ABB21653	Protein #3652 enco
5	8	3.7	159	AAW57060	Human brain expres
6	8	3.7	159	AAW69450	Human bone marrow
7	8	3.7	159	AAW17284	Peptide #3718 enco
8	8	3.7	159	AAW29784	Peptide #3821 enco
9	8	3.7	159	AAW4971	Peptide #3653 enco
10	8	3.7	159	ABG39070	Human peptide enco
C 11	8	3.7	342	AAW25835	Human platelet act
C 12	8	3.7	342	AAW49550	Human platelet act
C 13	8	3.7	342	ABB56381	Non-endogenous hum
C 14	8	3.7	342	ABB77904	Amino acid sequenc
C 15	8	3.7	342	ABB77918	Amino acid sequenc
16	8	3.7	485	AAW1002	Rat neuronal gluta
C 17	8	3.7	631	AAW56847	Arabidopsis thalia
C 18	8	3.7	634	AAW56846	Arabidopsis thalia
C 19	8	3.7	743	AAW56845	Arabidopsis thalia
20	8	3.7	2870	AAW95559	Caenorhabditis ele
21	8	3.7	3178	AAW95556	Caenorhabditis ele
22	7	3.2	15	ABG62591	Eubacterial DNA po
23	7	3.2	18	AAW92063	Modified type of e
24	7	3.2	18	AAW95416	Improved hen egg w
25	7	3.2	19	AAW74938	Human secreted pro
26	7	3.2	21	AAW80973	Sequence of modifi
27	7	3.2	21	AAW92065	N-terminal region
28	7	3.2	21	AAW05417	Signal peptide.
29	7	3.2	22	AAW00284	Human secreted pro
C 30	7	3.2	38	AAW53074	Human secreted pro
31	7	3.2	43	AAW57883	Arabidopsis thalia
32	7	3.2	44	AAW14432	Human secreted pro
33	7	3.2	46	AAW22521	Novel human colon
34	7	3.2	46	AAW92476	Human digestive sy
C 35	7	3.2	47	AAW02178	Human polypeptide
36	7	3.2	48	ABB95983	Human testicular a
37	7	3.2	48	AAW95286	Human reproductive
C 38	7	3.2	49	AAW06158	Arabidopsis thalia
C 39	7	3.2	52	AAW00502	Human secreted pro
40	7	3.2	55	AAW19543	Human acid sequenc
41	7	3.2	55	ABW04242	Human ORF protein
42	7	3.2	62	AAW10324	Pinto bean alpha-D
43	7	3.2	63	AAW07605	A human interleukin
44	7	3.2	63	AAW07696	A human interleukin
C 45	7	3.2	85	ABB70952	Drosophila melanog

ALIGNMENTS

RESULT 1
AAW40006
ID AAW40006 standard; peptide: 29 AA.
AC AAW40006;
XX 18-JUN-1998 (first entry)
DT Peptide effecting G-protein-coupled receptor activity.
XX
DE
XX
XX G-protein-coupled receptor; GPCR: transmembrane domain; oligomerisation;
KW therapeutic composition; GPCR function; receptor monomeric form;
KW multimeric form; inhibition; GPCR-mediated process; GPCR binding;
KW treatment; disease; platelet activating factor.
XX
XX Synthetic.
OS Homo sapiens.
OS
XX
PN W09800538-A2.

XX 08-JAN-1998.
PD
XX
XX 01-JUL-1997; 97WO-IB00814.
PF
XX
XX 01-JUL-1996; 96US-0021031.
PR
XX
PA (BIOS-) BIOSIGNAL INC.
PA (UYMO-) UNIV MONTREAL.
XX
PI Bouvier M, Dennis M, Hebert TE;
XX
DR WPI; 1998-086964/08.
XX
PT Peptide(s) or peptide leads affecting G protein-coupled receptor
PT activity - by altering receptor oligomerisation, useful in, e.g.
PT selectively modulating receptor function and treating neurological
PT or genetic diseases
XX
PS Claim 20; Page 62; 75pp; English.
XX
CC Peptides AAW4002-08 are modelled on transmembrane domains one to seven,
CC respectively, of human platelet activating factor, a G-protein-coupled
CC receptor (GPCR) (sic) whose activity is affected by the formation of
CC oligomers. GPCRs have a recurring pattern unique for the transmembrane
CC domains. The peptides are characterised by the ability to selectively
CC affect oligomerisation of the GPCR from which it was designed. The
CC peptides may be used with suitable carriers in the preparation of
CC therapeutic compositions. They may be administered to selectively
CC modulate GPCR function by affecting the ratio of receptor monomeric
CC to multimeric forms. The compositions can be administered to inhibit
CC GPCR-mediated processes by modulating GPCR binding to its agonist or
CC antagonist. The peptides may also be used to prevent or treat diseases
CC involving GPCRs. They are also useful for in vitro and in vivo studies
CC of GPCRs.
XX
SQ Sequence 29 AA;
XX
Alignment Scores:
Pred. No.: 52.2 Length: 29
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 19 Gaps: 0
US-09-835-992A-19 (1-714) x AAW4006 (1-29)
OY 117 TCTTTTCTAGTTTCTAATA 94
Db 11 SerpheeueuValPheleulle 18
RESULT 2
ABB31093
ID ABB31093 standard; Peptide; 159 AA.
XX
AC ABB31093;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #3744 encoded by breast cell single exon nucleic acid probe.
XX
KW Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
PF 30-JAN-2001; 2001WO-US00662.
XX

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
PS Claim 27; SEQ ID NO 14061; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BF 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 159 AA;
XX
Alignment Scores:
Pred. No.: 42.6 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
DB: 22 Gaps: 0
US-09-835-992A-19 (1-714) x ABB31093 (1-159)
OY 493 CTGTGGCATTGTTCTTACTT 516
Db 120 LeuLeuAlaLeuCySpheleulleu 127
RESULT 3
ABB36288
ID ABB36288 standard; Peptide; 159 AA.
XX
AC ABB36288;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #3794 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.

```
XX 30-JAN-2001; 2001WO-US00669.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
XX Claim 27; SEQ ID NO 28923; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 159 AA;

Alignment Scores:
Pred. No.: 42.6 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-19 (1-714) x ABB36288 (1-159)
OY 493 CTGTTGGCATTGTGTTCTTACTT 516
Db 120 LeuLeuAlaLeuCySpheLeuLeu 127

RESULT 4
ABB21653
ID ABB21653 standard; Protein; 159 AA.
XX
XX ABB21653;
AC
XX 23-JAN-2002 (first entry)
DT
XX
XX Protein #3652 encoded by probe for measuring heart cell gene expression.
DE
XX
XX Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
XX WO200157274-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00666.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
```

```
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX
XX Claim 15; SEQ ID NO 23423; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 159 AA;

Alignment Scores:
Pred. No.: 42.6 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-19 (1-714) x ABB21653 (1-159)
OY 493 CTGTTGGCATTGTGTTCTTACTT 516
Db 120 LeuLeuAlaLeuCySpheLeuLeu 127

RESULT 5
AAM57060
ID AAM57060 standard; Protein; 159 AA.
XX
XX AAM57060;
AC
XX 05-NOV-2001 (first entry)
DT
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 29165.
DE
XX
XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
XX WO200157275-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00667.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
```

PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
DR
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 29165; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 159 AA;

Alignment Scores:
Pred. No.: 42.6 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
DB: Gaps: 0

US-09-835-992A-19 (1-714) x AAM57060 (1-159)
QY 493 CTGTTGGCATGTGTTCTTACTT 516
XXXXXXXXXXXXXXXXXXXXXXXXXXXX
DB 120 LeuLeuAlaLeuCysPheLeuLeu 127

RESULT 6
AAM69450
ID AAM69450 standard; Protein; 159 AA.
XX
AC AAM69450;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 29756.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
DR
XX
PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 29756; 658pp + Sequence Listing; English.
PS
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 159 AA;

Alignment Scores:
Pred. No.: 42.6 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
DB: Gaps: 0

US-09-835-992A-19 (1-714) x AAM69450 (1-159)
QY 493 CTGTTGGCATGTGTTCTTACTT 516
XXXXXXXXXXXXXXXXXXXXXXXXXXXX
DB 120 LeuLeuAlaLeuCysPheLeuLeu 127

RESULT 7
AAM17284
ID AAM17284 standard; Protein; 159 AA.
XX
AC AAM17284;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #3718 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
DR
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 27; SEQ ID No 22110; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore

CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 159 AA;

Alignment Scores:

Pred. No.:	42.6	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	22	Gaps:	0

US-09-835-992A-19 (1-714) x AAM17284 (1-159)

OY 493 CTGTTGGCATTGTGTTCTTACTT 516

Db 120 LeuLeuAlaLeuCySpheLeuLeu 127

RESULT 8

AAM29784
ID AAM29784 standard; Protein; 159 AA.

XX
AC AAM29784;

DT 17-OCT-2001 (first entry)

DE Peptide #3821 encoded by probe for measuring placental gene expression.

XX
KW Probe; microarray; human; placenta; antenatal diagnosis;

XX
OS genetic disorder.

XX
PN Homo sapiens.

XX
WO200157272-A2.

XX
PD 09-AUG-2001.

XX
PF 30-JAN-2001; 2001WO-US00663.

XX
PR 04-FEB-2000; 2000US-0180312.

XX
PR 26-MAY-2000; 2000US-0207456.

XX
PR 30-JUN-2000; 2000US-0608408.

XX
PR 03-AUG-2000; 2000US-0632366.

XX
PR 21-SEP-2000; 2000US-0234687.

XX
PR 27-SEP-2000; 2000US-0236359.

XX
PR 04-OCT-2000; 2000GB-0024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
XX WPI; 2001-488897/53.

XX
PT Human genome-derived single exon nucleic acid probes useful for

XX
PS Claim 27; SEQ ID No 30053; 654pp; English.

XX
CC The present invention relates to single exon nucleic acid probes (SEND:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.

XX
SQ Sequence 159 AA;

Alignment Scores:

Pred. No.:	42.6	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	22	Gaps:	0

US-09-835-992A-19 (1-714) x AAM29784 (1-159)

OY 493 CTGTTGGCATTGTGTTCTTACTT 516

Db 120 LeuLeuAlaLeuCySpheLeuLeu 127

RESULT 9

AAM04971
ID AAM04971 standard; Protein; 159 AA.

XX
AC AAM04971;

DT 09-OCT-2001 (first entry)

DE Peptide #3653 encoded by probe for measuring breast gene expression.

XX
KW Probe; human; breast disease; breast cancer; development disorder;

XX
KW Inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX
OS Homo sapiens.

XX
PN WO200157270-A2.

XX
PD 09-AUG-2001.

XX
PF 29-JAN-2001; 2001WO-US00661.

XX
PR 04-FEB-2000; 2000US-0180312.

XX
PR 26-MAY-2000; 2000US-0207456.

XX
PR 30-JUN-2000; 2000US-0608408.

XX
PR 03-AUG-2000; 2000US-0632366.

XX
PR 21-SEP-2000; 2000US-0234687.

XX
PR 27-SEP-2000; 2000US-0236359.

XX
PR 04-OCT-2000; 2000GB-0024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
XX WPI; 2001-476286/51.

XX
PT Novel single exon nucleic acid probe used to measuring gene expression

XX
PS Claim 27; SEQ ID No 13711; 322pp; English.

XX
CC The present invention relates to novel single exon nucleic acid probes

XX
CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one

XX
CC such probe. The probes are useful for measuring human gene expression in

XX
CC a human breast sample, where the probe hybridises at high stringency to a

XX
CC nucleic acid expressed in the human breast. The probes are useful for

XX
CC predicting, diagnosing, grading, staging, monitoring and prognosing

XX
CC diseases of the human breast, particularly those diseases with polygenic

XX
CC aetiology. The diseases include: breast cancer, disorders of development,

XX
CC inflammatory diseases of the breast, fibrocystic changes, proliferative

XX
CC breast disease and non-carcinoma tumours.

XX
CC Note: The sequence data for this patent did not form part of the printed

XX
CC specification, but was obtained in electronic format directly from WIPO

XX
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 159 AA;

Alignment Scores:

Pred. No.:	42.6	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match: 3.67% Indels: 0
DB: 22 Gaps: 0
US-09-835-992A-19 (1-714) x AAM04971 (1-159)
OY 493 CTGTGGCATTGTTCTTACTT 516
Db 120 LeuLeuAlaLeucysPheLeuLeu 127
RESULT 10
ABG39070
ID ABG39070 standard; Peptide; 159 AA.
AC ABG39070;
DT 19-AUG-2002 (first entry)
DE Human peptide encoded by genome-derived single exon probe SEQ ID 28735.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PE 30-JAN-2001; 2001WO-US00665.
XX
PF 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
PS Claim 27; SEQ ID No 28735; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,

CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 159 AA;
XX
Alignment Scores:
Pred. No.: 42.6 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
DB: 23 Gaps: 0
US-09-835-992A-19 (1-714) x ABG39070 (1-159)
OY 493 CTGTGGCATTGTTCTTACTT 516
Db 120 LeuLeuAlaLeucysPheLeuLeu 127
RESULT 11
AAR25835
ID AAR25835 standard; Protein; 342 AA.
XX
AC AAR25835;
XX
DT 21-JAN-1993 (first entry)
DE Human platelet activating factor receptor.
XX
KW PAF; coagulation; blood clotting; clone phPAF.
XX
OS Homo sapiens.
XX
PN WO9212244-A.
XX
PD 23-JUL-1992.
XX
PF 27-DEC-1991; 91WO-JP01779.
XX
PR 27-DEC-1990; 90JP-0407119.
PR 27-DEC-1990; 90JP-0407945.
XX
PA (NISB) JAPAN TOBACCO INC.
XX
PI Honda Z, Nakamura M, Shimizu T;
XX
DR WPI; 1992-268667/32.
DR N-PSDB; AAQ26960.
XX
PT DNA coding for human or guinea pig platelet activating factor
PT receptor - for prodn. of recombinant PAF receptor for screening
PT potential PAF agonists-antagonists

XX Claim 3; Fig 5; 50pp; Japanese.
PS
XX
CC The cDNA coding for human PAF receptor was isolated in a 1780 bp
CC clone designated pHPAF. The clone consisted of a 112bp 5'-UTR (see
CC AAQ26961), a 1029bp ORF (AAQ26960) and a 639bp 3'-UTR (see AAQ26962).
CC The cDNA can be used to transform a suitable host cell line (e.g. COS
CC cells) for expression of the PAF receptor protein. The receptor
CC itself may be useful in screening of potential PAF agonists/
CC antagonists and in assays for PAF. See also AAQ26957-Q26959.
XX
SQ Sequence 342 AA;

Alignment Scores:
Pred. No.: 38.9 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 13 Gaps: 0

US-09-835-992A-19 (1-714) x AAR25835 (1-342)

QY 117 TCTTTTTCCTAGTTTCTAATA 94
Db 194 SerPhepheleuValpheleulle 201

RESULT 12
AAV49550
ID AAV49550 standard; Protein; 342 AA.
XX
AC AAV49550;
XX
DT 13-JAN-2000 (first entry)
XX
DE Human platelet activating factor receptor protein sequence.
XX
KW Human; coding sequence polymorphism; vascular pathology gene;
KW polymorphic site; phenotype correlation; forensic; paternity testing;
KW medicine; genetic analysis; vascular disease.
XX
OS Homo sapiens.
XX
PN WO950454-A2.
XX
PD 07-OCT-1999.
XX
PF 26-MAR-1999; 99WO-US06473.
XX
PR 01-APR-1998; 98US-0054272.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;
XX
DR WPI: 1999-620066/53.
DR N-PSDB: AAZ32159.
XX
PT Determination of polymorphisms in genes, especially those identifying
PT predisposition to vascular disease -
XX
PS Disclosure; Fig 3; 134pp; English.
XX
CC AAZ32159 to AAZ32194 represent reference alleles for specifically
CC claimed nucleic acid sequences from the present invention which comprise
CC polymorphic sites as given in a table in the specification, selected
CC from 92 single nucleotide polymorphisms in which the nucleotide at the
CC polymorphic site is different from a nucleotide at the same site in a
CC reference allele. The nucleic acids, and primers and probes, are used to
CC identify polymorphisms, which may predispose an individual to disease,
CC especially a vascular disease. They can also be used in phenotype
CC correlations, forensics, paternity testing, medicine or genetic
CC analysis. AAV49550 to AAV49573 represent the proteins which correspond

CC to some of the reference alleles.
XX
SQ Sequence 342 AA;

Alignment Scores:
Pred. No.: 38.9 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 20 Gaps: 0

US-09-835-992A-19 (1-714) x AAV49550 (1-342)

QY 117 TCTTTTTCCTAGTTTCTAATA 94
Db 194 SerPhepheleuValpheleulle 201

RESULT 13
ABB56381
ID ABB56381 standard; Protein; 342 AA.
XX
AC ABB56381;
XX
DT 18-FEB-2002 (first entry)
XX
DE Non-endogenous human GPCR protein, SEQ ID NO: 555.
XX
KW Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
KW constitutively activated GPCR; agonist; disease.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN WO200177172-A2.
XX
PD 18-OCT-2001.
XX
PF 05-APR-2001; 2001WO-US11098.
XX
PR 07-APR-2000; 2000US-195747P.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Lehmann-Brulnsma K, Liaw CW, Lhn I;
XX
DR WPI: 2001-648759/74.
DR N-PSDB; ABI98017.
XX
PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in
PT disease treatment, comprises contacting candidate compounds with
PT versions of GPCRs -
XX
PS Claim 1; Page 356-357; 394pp; English.
XX
CC The invention relates to G protein-coupled receptors (GPCRs) for which
CC the endogenous ligand has been identified. Non-endogenous
CC constitutively activated versions of known GPCRs are used in the
CC invention for the direct identification of candidate compounds as
CC receptor agonists, inverse agonists or partial agonists. Such
CC agonists are useful as therapeutic agents for diseases or disorders
CC associated with GPCRs. The present sequence is a non-endogenous
CC version of a known human GPCR.
XX
SQ Sequence 342 AA;

Alignment Scores:
Pred. No.: 38.9 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-19 (1-714) x ABB56381 (1-342)

OY 117 TCTTTTCCAGTTTCTAATA 94
Db 194 SerpheeValPheLeuIle 201

RESULT 14
ABB77904
ID ABB77904 standard; Protein; 342 AA.

AC ABB77904;

DT 07-OCT-2002 (first entry)

DE Amino acid sequence of platelet activating factor receptor.

KW Human; platelet activating factor receptor; PTAFR; gene; isogene;
KW chromosome 1; inflammatory disease; coronary disease; cancer;
KW receptor.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Misc-difference 28 /note= "when PS2 is T, this residue is phe"

FT Misc-difference 224 /note= "when PS3 is A, this residue is Asp"

FT Misc-difference 338 /note= "when PS5 is G, this residue is Ser"

PN WO200251859-A2.

PD 04-JUL-2002.

PF 05-NOV-2001; 2001WO-US47441.

PR 03-NOV-2000; 2000US-245633P.

PA (GENA-) GENAISSANCE PHARM INC.

PI Chew A, Choi JY, Koshy B;

DR WPI; 2002-566672/60.

DR N-PSDB; ABL59299.

PT New genetic variants comprising haplotypes of the human Platelet
PT Activating Factor Receptor (PTAFR) gene, useful for treating or
PT screening drugs for treating e.g. inflammatory diseases, coronary
PT diseases or cancer -

PS Claim 29; Fig 3; 59pp; English.

CC The present sequence represents a human Platelet Activating Factor
CC Receptor (PTAFR). The PTAFR gene comprises polymorphic sites referred to
CC as PS1-5 to designate the order in which they are located in the gene.
CC Six isogenes of the PTAFR gene exist. The PTAFR gene is located on
CC chromosome 1, and contains 1 exon. Polymorphisms PS3 and PS5 have
CC previously been identified. PS3 and PS5 occur in the coding region.
CC The polynucleotide comprising polymorphisms in the PTAFR gene is
CC useful in screening candidate drugs to treat diseases related to PTAFR
CC activity, e.g. inflammatory diseases, coronary diseases or cancer. The
CC PTAFR isogenes are especially useful for treating these diseases. The
CC methods and haplotypes are useful in improving the efficiency of drug
CC discovery and development processes, or for designing clinical trials
CC of candidate drugs for treating the specific condition or disease
CC described above.

SQ Sequence 342 AA;

Alignment Scores:

Pred. No.: 38.9 Length: 342
Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 23 Gaps: 0

US-09-835-992A-19 (1-714) x ABB77904 (1-342)

OY 117 TCTTTTCCAGTTTCTAATA 94
Db 194 SerpheeValPheLeuIle 201

RESULT 15
ABB77918
ID ABB77918 standard; Protein; 342 AA.

AC ABB77918;

DT 07-OCT-2002 (first entry)

DE Amino acid sequence of platelet activating factor receptor variant.

KW Human; platelet activating factor receptor; PTAFR; gene; isogene;
KW chromosome 1; inflammatory disease; coronary disease; cancer;
KW receptor.

OS Homo sapiens.

PN WO200251859-A2.

PD 04-JUL-2002.

PF 05-NOV-2001; 2001WO-US47441.

PR 03-NOV-2000; 2000US-245633P.

PA (GENA-) GENAISSANCE PHARM INC.

PI Chew A, Choi JY, Koshy B;

DR WPI; 2002-566672/60.

PT New genetic variants comprising haplotypes of the human Platelet
PT Activating Factor Receptor (PTAFR) gene, useful for treating or
PT screening drugs for treating e.g. inflammatory diseases, coronary
PT diseases or cancer -

PS Claim 32; Page -: 59pp; English.

CC The present sequence represents a human Platelet Activating Factor
CC Receptor (PTAFR) polymorphic variant. The PTAFR gene has polymorphic
CC sites referred to as PS1-5 to designate the order in which they are
CC located in the gene. Six isogenes of the PTAFR gene exist. The PTAFR
CC gene is located on chromosome 1, and contains 1 exon. Polymorphisms PS3
CC and PS5 have previously been identified. PS3 and PS5 occur in the coding
CC region. The polynucleotide comprising polymorphisms in the PTAFR gene is
CC useful in screening candidate drugs to treat diseases related to PTAFR
CC activity, e.g. inflammatory diseases, coronary diseases or cancer. The
CC PTAFR isogenes are especially useful for treating these diseases. The
CC methods and haplotypes are useful in improving the efficiency of drug
CC discovery and development processes, or for designing clinical trials
CC of candidate drugs for treating the specific condition or disease
CC described above.
CC note: this sequence does not appear in the specification; it was created
CC using information provided.

SQ Sequence 342 AA;

Alignment Scores:

Pred. No.: 38.9 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0

Wed Jan 15 08:40:57 2003

DB:	23	Gaps:	0
US-09-835-992A-19 (1-714) x ABB77918 (1-342)			
OY	117 TCTTTTCTAGTTTCTATA 94		
Db	194 SerPheLeuValPheLeuIle 201		

Search completed: January 14, 2003, 17:27:17
Job time : 38.1708 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:29:25 ; Search time 6.8357 Seconds
(without alignments)
4052.943 Million cell updates/sec

Title: US-09-835-992A-19
Perfect score: 218
Sequence: 1 cgcacagaaaaagtatttta.....cacagtggttgccaattca 714

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 118974 seqs, 19401057 residues

Word size: 1

Total number of hits satisfying chosen parameters: 214280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+n2p.model -DEV-xlp
-O-/cgn2_1/USPTO_spool/US09835992/runat_14012003_161528_4118/app_query.fasta_1.3932
-DB-Published_Applications_AA -QFMT-fastan -SUFFIX-olig.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-oligo
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER-US09835992_@CGN_1_1_28_@runat_14012003_161528_4118
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCF_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCFUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	8	3.7	159 10	US-09-864-761-36951
2	8	3.7	327 12	US-10-084-206-3
3	7	3.2	81 9	US-10-002-344A-243
4	7	3.2	127 10	US-09-764-877-1466

5	7	3.2	197 10	US-09-780-717-47	Sequence 47, Appl
6	7	3.2	241 10	US-09-740-668A-26	Sequence 26, Appl
7	7	3.2	267 10	US-09-908-855-36	Sequence 36, Appl
8	7	3.2	334 10	US-09-815-242-13249	Sequence 13249, A
9	7	3.2	380 10	US-09-205-658-110	Sequence 110, Appl
10	7	3.2	380 10	US-09-844-353A-110	Sequence 110, Appl
11	7	3.2	420 10	US-09-812-862-10	Sequence 10, Appl
12	7	3.2	481 10	US-09-815-242-10888	Sequence 10888, A
13	7	3.2	554 10	US-09-800-396-4	Sequence 4, Appl
14	7	3.2	1724 10	US-09-205-658-12	Sequence 12, Appl
15	7	3.2	1724 10	US-09-844-353A-12	Sequence 12, Appl
16	6	2.8	9 10	US-09-834-765-444	Sequence 444, Appl
17	6	2.8	10 8	US-08-452-843A-13	Sequence 13, Appl
18	6	2.8	10 10	US-09-834-765-82	Sequence 82, Appl
19	6	2.8	10 10	US-09-834-765-511	Sequence 611, Appl
20	6	2.8	11 9	US-09-909-460-67	Sequence 67, Appl
21	6	2.8	17 10	US-09-949-196-4	Sequence 4, Appl
22	6	2.8	22 10	US-09-949-196-32	Sequence 32, Appl
23	6	2.8	22 10	US-09-949-196-36	Sequence 36, Appl
24	6	2.8	22 10	US-09-949-196-40	Sequence 40, Appl
25	6	2.8	25 12	US-10-001-879-113	Sequence 113, Appl
26	6	2.8	28 10	US-09-925-299-1023	Sequence 1023, Appl
27	6	2.8	28 10	US-09-929-818-104	Sequence 104, Appl
28	6	2.8	30 10	US-09-071-838-227	Sequence 227, Appl
29	6	2.8	31 10	US-09-864-761-38920	Sequence 38920, A
30	6	2.8	32 10	US-09-864-761-47089	Sequence 47089, A
31	6	2.8	34 10	US-09-864-761-34176	Sequence 34176, A
32	6	2.8	36 10	US-09-864-761-47949	Sequence 47949, A
33	6	2.8	36 10	US-09-864-761-48325	Sequence 48325, A
34	6	2.8	37 10	US-09-864-761-44137	Sequence 44137, A
35	6	2.8	38 10	US-09-864-761-46306	Sequence 46306, A
36	6	2.8	39 10	US-09-764-847-937	Sequence 937, Appl
37	6	2.8	40 10	US-09-864-761-41296	Sequence 41296, A
38	6	2.8	40 10	US-09-864-761-42251	Sequence 42251, A
39	6	2.8	41 10	US-09-864-761-46409	Sequence 46409, A
40	6	2.8	43 10	US-09-864-761-44190	Sequence 44190, A
41	6	2.8	43 10	US-09-764-847-866	Sequence 866, Appl
42	6	2.8	44 10	US-09-864-761-47765	Sequence 47765, A
43	6	2.8	45 10	US-09-764-877-1815	Sequence 1815, Appl
44	6	2.8	46 10	US-09-764-877-157	Sequence 157, Appl
45	6	2.8	47 10	US-09-925-299-785	Sequence 785, Appl

ALIGNMENTS

RESULT 1
US-09-864-761-36951
Sequence 36951, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Hanzel, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36951
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
;
; FEATURE:
; OTHER INFORMATION: MAP TO AC007374.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: AA485663.1, EVALUATE 1.00e-67
;
US-09-864-761-36951

Alignment Scores:
Pred. No.: 5.88 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-19 (1-714) x US-09-864-761-36951 (1-159)

QY 493 CTGTGGCATGTGTTCTTACTT 516
Db 120 LeuLeuAlaLeuCysPheLeuLeu 127

RESULT 2
US-10-084-206-3
; Sequence 3, Application US/10084206
; Patent No. US20020106741A1
; GENERAL INFORMATION:
; APPLICANT: Li et al.
; TITLE OF INVENTION: G-Protein Receptor HTNAD29
; FILE REFERENCE: P191D1C1
; CURRENT APPLICATION NUMBER: US/10/084,206
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US95/07288
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/468,534
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 09/399,095
; PRIOR FILING DATE: 1999-09-20
```

```
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Version 3.1
; SEQ ID NO 3
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-10-084-206-3

Alignment Scores:
Pred. No.: 5.21 Length: 327
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 12 Gaps: 0

US-09-835-992a-19 (1-714) x US-10-084-206-3 (1-327)

QY 117 TCTTTTTCCTAGTTTCTATA 94
Db 190 SerPhePheLeuValPheLeuIle 197

RESULT 3
US-10-002-344A-243
; Sequence 243, Application US/10002344A
; Patent No. US20020172959A1
; GENERAL INFORMATION:
; APPLICANT: Recipon, Herve
; APPLICANT: Sun, Yongming
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and P
; FILE REFERENCE: DEX-0241
; CURRENT APPLICATION NUMBER: US/10/002,344A
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/242,998
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 243
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-10-002-344A-243

Alignment Scores:
Pred. No.: 61.2 Length: 81
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 9 Gaps: 0

US-09-835-992a-19 (1-714) x US-10-002-344A-243 (1-81)

QY 621 ATTTTACTGTTCTTAATTTA 641
Db 73 IleLeuLeuPheLeuIleLeu 79

RESULT 4
US-09-764-877-1466
; Sequence 1466, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
```


; SEQ ID NO 1466
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1466

Alignment Scores:
Pred. No.: 56.7 Length: 127
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-764-877-1466 (1-127)

OY 326 GCCTGCTCAAGCTCTGATC 346
Db 80 AlAcysvalGlnAlaLeuile 86

RESULT 5
US-09-780-717-47
; Sequence 47, Application US/09780717
; Patent NO. US20010044941A1
; GENERAL INFORMATION:
; APPLICANT: Helentjaris, Tim
; APPLICANT: Bates, Nic
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: No. US20010044941A1el Invertase Inhibitors and Methods
; FILE REFERENCE: 035718/208677
; CURRENT APPLICATION NUMBER: US/09/780,717
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,509
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Glycine max
US-09-780-717-47

Alignment Scores:
Pred. No.: 52.6 Length: 197
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-780-717-47 (1-197)

OY 43 CAAGCATTTATTTATCCCTAT 63
Db 45 GlncserllelleuSerTyr 51

RESULT 6
US-09-740-668A-26
; Sequence 26, Application US/09740668A
; Patent NO. US20020076700A1
; GENERAL INFORMATION:
; APPLICANT: Shlmkels, Richard
; TITLE OF INVENTION: No. US20020076700A1el polypeptides and nucleic acids encoding sam
; FILE REFERENCE: 15966-537 CIP
; CURRENT APPLICATION NUMBER: US/09/740,668A
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: PCT/US99/29584
; PRIOR FILING DATE: 1999-12-17

; PRIOR APPLICATION NUMBER: 09/465,512
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113,485
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/112,837
; PRIOR FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)..(241)
; OTHER INFORMATION: wherein Xaa is any amino acid as defined in the specification.
US-09-740-668A-26

Alignment Scores:
Pred. No.: 50.9 Length: 241
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-740-668A-26 (1-241)

OY 338 GCCTGATCATATTTCTTTTA 358
Db 126 AlaLeuilellepheLeu 132

RESULT 7
US-09-908-855-36
; Sequence 36, Application US/09908855
; Patent NO. US20020115220A1
; GENERAL INFORMATION:
; APPLICANT: KONDO, KEIJI
; APPLICANT: MIURA, YUTAKA
; TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE
; FILE REFERENCE: 049441/0118
; CURRENT APPLICATION NUMBER: US/09/908,855
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 09/242,690
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: JP 8/241062
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Candida utilis
US-09-908-855-36

Alignment Scores:
Pred. No.: 50 Length: 267
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-908-855-36 (1-267)

OY 521 GAAGTAAGTAAGAACACAT 501
Db 80 GlncSerllyslsHisasn 86

RESULT 8
US-09-815-242-13249
; Sequence 13249, Application US/09815242


```
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13249
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-815-242-13249

Alignment Scores:
Pred. No.: 48.1 Length: 334
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-815-242-13249 (1-334)

QY 402 GCATTTCATGCTATCAGAAAC 382
Db 121 AlaphehisAlaileAArgAsn 127

RESULT 9
US-09-205-658-110
; Sequence 110, Application US/09205658
; Patent No. US20010029617A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; CURRENT APPLICATION NUMBER: US/09/205,658
; PRIOR FILING DATE: 1998-12-03
; PRIOR APPLICATION NUMBER: 08/857,076
; PRIOR FILING DATE: 1997-05-15
; PRIOR APPLICATION NUMBER: 08/888,534
; PRIOR FILING DATE: 1997-07-07
; PRIOR APPLICATION NUMBER: US98/10080
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 380
```

```
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-205-658-110

Alignment Scores:
Pred. No.: 47.1 Length: 380
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-205-658-110 (1-380)

QY 108 CTAGTTTTCCTAATCTAATG 88
Db 4 LeuValPheLeuIleLeuMet 10

RESULT 10
US-09-844-353A-110
; Sequence 110, Application US/09844353A
; Patent No. US20020037585A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutaru
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Kowek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; CURRENT APPLICATION NUMBER: US/09/844,353A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 08/857,076
; PRIOR FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-09-844-353A-110

Alignment Scores:
Pred. No.: 47.1 Length: 380
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-844-353A-110 (1-380)

QY 108 CTAGTTTTCCTAATCTAATG 88
Db 4 LeuValPheLeuIleLeuMet 10

RESULT 11
US-09-812-862-10
; Sequence 10, Application US/09812862
; Patent No. US20020035081A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: Scaglioni, Pier Paolo
; APPLICANT: Melegari, Margherita
; TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
```

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/812,862
FILING DATE: 20-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,073
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/492,489
FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/282001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 420 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-812-862-10
Alignment Scores:
Pred. No.: 46.3 Length: 420
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 10 Gaps: 0
US-09-835-992a-19 (1-714) x US-09-812-862-10 (1-420)
QY 514 CTTAGTTCTCCCAAGCAAAA 534
Db 292 LeuSerSerProLysGlyLys 298
RESULT 12
US-09-815-242-10888
Sequence 10888, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10888
LENGTH: 481
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10888
Alignment Scores:
Pred. No.: 45.3 Length: 481
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 10 Gaps: 0
US-09-835-992a-19 (1-714) x US-09-815-242-10888 (1-481)
QY 205 AAACGTCATCTGAAGCAAA 225
Db 399 LysThrAlaSerGlySerLys 405
RESULT 13
US-09-800-396-4
Sequence 4, Application US/09800396
Patent No. US20020065395A1
GENERAL INFORMATION:
APPLICANT: Wallis, Nicola
TITLE OF INVENTION: NOVEL RESPONSE REGULATOR
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/800,396
FILING DATE: 06-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/879,531
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER: GM10018
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 554 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-800-396-4

Alignment Scores:

Pred. No.: 44.2 Length: 554
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-800-396-4 (1-554)

OY 111 TTCCTAGTTTCTAATACTA 91

Db 6 PheValPheLeuIleLeu 12

RESULT 14

US-09-205-658-12

; Sequence 12, Application US/09205658

; Patent No. US20010029617A1

; GENERAL INFORMATION:

; APPLICANT: Ruvkun, Gary

; APPLICANT: Ogg, Scott

; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR

; FILE REFERENCE: 00786/351004

; CURRENT APPLICATION NUMBER: US/09/205,658

; EARLIER FILING DATE: 1998-12-03

; EARLIER APPLICATION NUMBER: 08/857,076

; EARLIER FILING DATE: 1997-05-15

; EARLIER APPLICATION NUMBER: 08/888,534

; EARLIER FILING DATE: 1997-07-07

; EARLIER APPLICATION NUMBER: US98/10080

; NUMBER OF SEQ ID NOS: 328

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 1724

; TYPE: PRT

; ORGANISM: Caenorhabditis elegans

US-09-205-658-12

Alignment Scores:

Pred. No.: 36.5 Length: 1724
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-205-658-12 (1-1724)

OY 108 CTAGTTTCTAATACTAATG 88

Db 1068 LeuValPheLeuIleLeuMet 1074

RESULT 15

US-09-844-353A-12

; Sequence 12, Application US/09844353A

; Patent No. US20020037585A1

; GENERAL INFORMATION:

; APPLICANT: Ruvkun, Gary

; APPLICANT: Kimura, Koutarou

; APPLICANT: Patterson, Garth

; APPLICANT: Ogg, Scott

; APPLICANT: Paradis, Suzanne

; APPLICANT: Tissenbaum, Heidi

; APPLICANT: Morris, Jason

; APPLICANT: Kowek, Allison

; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR

; FILE REFERENCE: 00786/351005

; CURRENT APPLICATION NUMBER: US/09/844,353A

; CURRENT FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 08/857,076

; PRIOR FILING DATE: 1997-05-15

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 1724

; TYPE: PRT

; ORGANISM: Caenorhabditis elegans

US-09-844-353A-12

Alignment Scores:

Pred. No.: 36.5 Length: 1724
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-844-353A-12 (1-1724)

OY 108 CTAGTTTCTAATACTAATG 88

Db 1068 LeuValPheLeuIleLeuMet 1074

Search completed: January 14, 2003, 18:08:15
Job time : 10.8357 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:45 ; Search time 23.925 Seconds
(without alignments)
5737.940 Million cell updates/sec

Title: US-09-835-992A-19
Perfect score: 218
Sequence: 1 cgcacagaaaaagatatttta.....cacagttgttgccaattca 714

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565536

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODEL-frame+n2p.model -DEV-xlp
-O-/cgn2.1/USPTO_spool/US09835992/runat_14012003_161526_4038/app_query_fasta_1.3932
-DB-PIR_73 -QFMT-fastan -SUFFIX-oligo.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.cdl -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE-LOCAL -OUTFMT-ptc
-NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER-US09835992_@CGN_1_1_192_@runat_14012003_161526_4038 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.7	195	2	E71000
2	8	3.7	342	2	A40191
3	8	3.7	343	2	T19082
4	8	3.7	393	2	H64233
5	8	3.7	512	2	E83060
6	8	3.7	791	2	H72552
7	8	3.7	1118	1	A49724
8	8	3.7	1183	2	F90559
9	8	3.7	2329	2	T28125
10	8	3.2	65	2	S34538
11	7	3.2	86	2	S21595
12	7	3.2	105	2	E84075
13	7	3.2	113	2	C70342
14	7	3.2	120	2	T30769

15	7	3.2	126	2	F72150	B6L protein - var1
16	7	3.2	164	1	NNMU1	2S albumin 1 precu
17	7	3.2	164	1	NNMU3	2S albumin 3 precu
18	7	3.2	194	2	C69933	negative regulatio
19	7	3.2	201	2	G90674	probable transcrip
20	7	3.2	201	2	B85525	probable transcrip
21	7	3.2	201	2	F64758	yahd protein - Esc
22	7	3.2	205	2	T08398	calmodulin homolog
23	7	3.2	207	2	T24267	hypothetical prote
24	7	3.2	208	1	A63113	conserved hypotet
25	7	3.2	214	2	G70348	hypothetical prote
26	7	3.2	236	1	B64212	probable integral
27	7	3.2	237	2	C81291	probable integral
28	7	3.2	245	2	A30154	IgE receptor alpha
29	7	3.2	249	2	C81449	probable molybdate
30	7	3.2	254	2	D69140	hypothetical prote
31	7	3.2	271	2	F72379	sugar ABC transpor
32	7	3.2	278	2	B75442	3-hydroxybutyryl-C
33	7	3.2	279	2	T20443	hypothetical prote
34	7	3.2	290	2	G72203	sugar ABC transpor
35	7	3.2	307	2	G64350	quinolinate synthe
36	7	3.2	314	2	T45010	probable site-spec
37	7	3.2	314	2	S73717	hypothetical prote
38	7	3.2	314	2	E84970	pseudouridylylate sy
39	7	3.2	317	2	B71110	hypothetical prote
40	7	3.2	319	2	D90589	lipase homolog T0
41	7	3.2	320	2	T30894	glycoprotein VP7 p
42	7	3.2	326	1	VGXRHH	glycoprotein VP7 p
43	7	3.2	326	1	VGXRHR	glycoprotein VP7 p
44	7	3.2	326	1	A44891	conserved hypotet
45	7	3.2	334	2	B98019	

ALIGNMENTS

RESULT 1
E71000
hypothetical protein PH1301 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 18-Aug-2000
C:Accession: E71000
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: E71000
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-195 <KAW>
A:Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30405.1; PID:d1031348; PID:g32
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBa
C:Genetics:
A:Gene: PH1301
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1301

Alignment Scores:
Pred. No.: 10.6
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.69%
DB: 2
Gaps: 0

US-09-835-992A-19 (1-714) x E71000 (1-195)

QY 120 TTAATCTTTTCTAGTTTCTA 97
Db 99 LeuSerPheLeuValPheLeu 106

RESULT 2
A40191

platelet-activating factor receptor - human
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 20-Jun-2000
C:Accession: A40191; JH0479; A41079; JCI359; A42831; I51923
R:Kunz, D.; Gerard, N.P.; Gerard, C.
J. Biol. Chem. 267, 9101-9106, 1992
A:Title: The human leukocyte platelet-activating factor receptor. cDNA cloning, cell su
A:Reference number: A40191; MUID:92250505; PMID:1374385
A:Accession: A40191
A:Molecule type: mRNA
A:Residues: 1-342 <KUN>
A:Cross-references: GB:M76674; NID:g456293; PIDN:AAA60002.1; PID:g456294
R:Ye, R.D.; Prossnitz, E.R.; Zou, A.; Cochrane, C.G.
Biochem. Biophys. Res. Commun. 180, 105-111, 1991
A:Title: Characterization of a human cDNA that encodes a functional receptor for platele
A:Reference number: JH0479; MUID:92028922; PMID:1656963
A:Accession: JH0479
A:Molecule type: mRNA
A:Residues: 1-342 <YER>
A:Cross-references: GB:M80436; NID:g189537; PIDN:AAA60001.1; PID:g189538
A:Experimental source: granulocyte, cell line HL-60 all
R:Nakamura, M.; Honda, Z.; Izumi, T.; Sakanaoka, C.; Mutoh, H.; Minami, M.; Bito, H.; Sey
J. Biol. Chem. 266, 20400-20405, 1991
A:Title: Molecular cloning and expression of platelet-activating factor receptor from hu
A:Reference number: A41079; MUID:92041873; PMID:1657923
A:Accession: A41079
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-342 <NAK>
A:Cross-references: GB:D10202; GB:D90433; NID:g219975; PIDN:BAA01050.1; PID:g219976
R:Sugimoto, T.; Tsuchimochi, H.; McGregor, C.G.A.; Mutoh, H.; Shimizu, T.; Kurechl, Y.
Biochem. Biophys. Res. Commun. 189, 617-624, 1992
A:Title: Molecular cloning and characterization of the platelet-activating factor recept
A:Reference number: JCI359; MUID:93112021; PMID:1281995
A:Accession: JCI359
A:Molecule type: mRNA
A:Residues: 1-315, 'N', 317-342 <SUG>
A:Experimental source: heart
A:Note: the authors translated the codon AAT for residue 316 as Lys
R:Seyfried, C.E.; Schweickart, V.L.; Godiska, R.; Gray, P.W.
Genomics 13, 832-834, 1992
A:Title: The human platelet-activating factor receptor gene (PTAFR) contains no introns
A:Reference number: A42831; MUID:92347886; PMID:1322356
A:Accession: A42831
A:Molecule type: DNA
A:Residues: 1-226, 'TG', 229-342 <SEY>
A:Cross-references: GB:M88177; NID:g190697; PIDN:AAA60214.1; PID:g190698
A:Note: sequence extracted from NCBI backbone (NCBIN:109813, NCBIP:109814)
R:Chase, P.B.; Halonen, M.; Regan, J.W.
Am. J. Respir. Cell Mol. Biol. 8, 240-244, 1993
A:Title: Cloning of a human platelet-activating factor receptor gene: evidence for an in
A:Reference number: I51923; MUID:93192035; PMID:8383507
A:Accession: I51923
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-342 <RES>
A:Cross-references: GB:S56396; NID:g298580; PIDN:AAB25755.1; PID:g298581
C:Genetics:
A:Gene: GDB:PTAFR
A:Cross-references: GDB:128806; OMIM:173393
A:Map position: 1p35-1p34.3
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; transmembrane protein
F:17-38/Domain: transmembrane #status predicted <TRI>
F:54-75/Domain: transmembrane #status predicted <TII>
F:92-113/Domain: transmembrane #status predicted <III>
F:134-155/Domain: transmembrane #status predicted <TIV>
F:184-205/Domain: transmembrane #status predicted <TRV>
F:233-253/Domain: transmembrane #status predicted <TVI>
F:277-297/Domain: transmembrane #status predicted <VII>
Alignment Scores:
Pred. No.: 10 Length: 342

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 2 Gaps: 0
US-09-835-992A-19 (1-714) x A40191 (1-342)
OY 117 TCCTTTTCCTAGTTTCTAATA 94
|||||
Db 194 SerpHeLeuValPheLeuIle 201
RESULT 3
T19082
hypothetical protein C08B6.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19082
R:Wilkinson, J.
submitted to the EMBL Data Library, May 1996
A:Reference number: T19070
A:Accession: T19082
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-343 <WIL>
A:Cross-references: EMBL:Z72502; PIDN:CAA96591.1; GSPDB:GN00023; CESP:C08B6.9
A:Experimental source: clone C08B6
C:Genetics:
A:Gene: CESP:C08B6.9
A:Map position: 5
A:Introns: 25/2; 84/3; 174/2; 290/2
Alignment Scores:
Pred. No.: 10 Length: 343
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 2 Gaps: 0
US-09-835-992A-19 (1-714) x T19082 (1-343)
OY 646 CAGATAAAATTGACACGTAAA 623
|||||
Db 23 GlnAsnLysIleArgAsnSerLys 30
RESULT 4
H64233
hypothetical protein MG306 - Mycoplasma genitalium
C:Species: Mycoplasma genitalium
C>Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 07-Dec-1999
C:Accession: H64233
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,
C.A.; Venter, J.C.
Science 270, 397-403, 1995
A:Title: The minimal gene complement of Mycoplasma genitalium.
A:Reference number: A64200; MUID:96026346; PMID:7569993
A:Accession: H64233
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-393 <TIGR>
A:Cross-references: GB:U39712; GB:L43967; NID:g1046005; PID:g1046006; TIGR:MG306
A:Experimental source: strain G-37
C:Genetics:
A:Genetic code: SGC3
Alignment Scores:
Pred. No.: 9.88 Length: 393
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0

DB: 2 Gaps: 0

US-09-835-992A-19 (1-714) x H64233 (1-393)

OY 117 TCTTTTTCCTAGTTTCTAATA 94
|||||
Db 19 SerpHeLeuValPheLeuIle 26

RESULT 5

E83060

Iron (III)-transport system permease HltB PA4688 [imported] - Pseudomonas aeruginosa (strain ATCC 27803)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: E83060

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim, J.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: E83060

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-512 <STO>

A:Cross-references: GB:AE004882; GB:AE004091; NID:9950939; PIDN:AG08075.1; GSPDB:GN001

C:Experimental source: strain PA01

C:Genetics:

A:Gene: hltB; PA4688

C:Superfamily: sfub protein

Alignment Scores:

Pred. No.:	9.61	Length:	512
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-19 (1-714) x E83060 (1-512)

OY 493 CTGTTGGCATTGTGTTCTTACTT 516
|||||
Db 230 LeuLeuAlaLeuCysPheLeu 237

RESULT 6

H72552

hypothetical protein APE1708 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: H72552

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Nakawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kawauchi, Y.

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: H72552

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-791 <KAW>

A:Cross-references: DDBJ:AP000062; NID:95105244; PIDN:BA080709.1; PID:dl044495; PID:95105244

C:Experimental source: strain K1

C:Genetics:

A:Gene: APE1708

Alignment Scores:

Pred. No.:	9.19	Length:	791
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-19 (1-714) x H72552 (1-791)

OY 90 TTACTATTAGAAAACTAGAAA 113
|||||
Db 134 LeuValLeuGluLysLeuGlyLys 141

RESULT 7

A49724

protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type II precursor - human

N:Alternate names: protein-tyrosine-phosphatase, stomach cancer-associated type 1; SA

C:Species: Homo sapiens (man)

C>Date: 03-May-1994 #sequence_revision 23-Feb-1996 #text_change 16-Jun-2000

C:Accession: A49724

R:Matozaki, T.; Suzuki, T.; Uchida, T.; Inazawa, J.; Ariyama, T.; Matsuda, K.; Horita, J. Biol. Chem. 269, 2075-2081, 1994

A:Title: Molecular cloning of a human transmembrane-type protein tyrosine phosphatase

A:Reference number: A49724; MUID:94124561; PMID:8294459

A:Accession: A49724

A:Molecule type: mRNA

A:Residues: 1-1118 <MATO>

A:Cross-references: GB:D15049; NID:9475003; PIDN:BA03645.1; PID:9475004

C:Genetics:

A:Gene: GDB:PTPRH; SAP-1

A:Cross-references: GDB:305504

A:Map position: 19q13.4-19q13.4

A>Note: highly expressed in colon and pancreatic cancer cells but not in the normal cells

C:Superfamily: protein-tyrosine-phosphatase, receptor type II; fibronectin type III repeat

C:Keywords: carcinogenesis; duplication; glycoprotein; phosphoprotein; phosphoric monophosphate

F:1-27/Domain: signal sequence #status predicted <SIG>

F:27-110/Domain: fibronectin type III repeat homology <3FNA>

F:28-118/Domain: protein-tyrosine-phosphatase, receptor type II #status predicted <M>

F:28-761/Domain: extracellular #status predicted <EXT>

F:116-199/Domain: fibronectin type III repeat homology <3FNB>

F:205-289/Domain: fibronectin type III repeat homology <3FNC>

F:296-379/Domain: fibronectin type III repeat homology <3FND>

F:385-468/Domain: fibronectin type III repeat homology <3FNE>

F:474-558/Domain: fibronectin type III repeat homology <3FNF>

F:564-658/Domain: fibronectin type III repeat homology <3FNG>

F:667-737/Domain: fibronectin type III repeat homology <3FNH>

F:762-778/Domain: transmembrane #status predicted <TMN>

F:779-1118/Domain: intracellular #status predicted <INT>

F:846-1070/Domain: protein-tyrosine-phosphatase homology <PTP1>

F:35,78,83,107,132,149,172,196,203,286,304,312,329,352,376,383,401,436,439,470,490,555

F:1022/Active site: Cys (phosphocysteine intermediate) #status predicted

F:1028/Binding site: substrate phosphate (Arg) #status predicted

Alignment Scores:

Pred. No.:	8.86	Length:	1118
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	1	Gaps:	0

US-09-835-992A-19 (1-714) x A49724 (1-1118)

OY 618 GGGATTCTACTGTCTCAATTTA 641
|||||
Db 763 GlyIleLeuLeuPheLeuIleLeu 770

RESULT 8

F90559

conserved hypothetical protein MYPU_3820 [imported] - Mycoplasma pulmonis (strain UAB)

C:Species: Mycoplasma pulmonis

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001

C:Accession: F90559

R:Chambers, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, J.

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis

A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: F90559

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1183 <KUR>

A:Cross-references: GB:AL445566; PID:g14089796; PIDN:CAC13555.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV_3820
A:Genetic code: SGC3

Alignment Scores:

Pred. No.:	8.81	Length:	1183
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-19 (1-714) x F90559 (1-1183)

OY 208 TTTTCAATTAATCTCAATTCA 185
Db 577 PheSerAsnTyrAsnLeuIleSer 584

RESULT 9

T28125

hypothetical protein ZK945.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T28125

R:Wilkinson, J.

A:Reference number: Z20472

A:Accession: T28125

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2329 <WIL>

A:Cross-references: EMBL:Z48544; NID:g695493; PIDN:CAA88442.1; GSPDB:GN00020; CESP:ZK945

A:Experimental source: clone ZK945

C:Genetics:

A:Gene: CESP:ZK945.9

A:Map position: 2

A:Introns: 160/3; 187/1; 243/1; 580/3; 607/3; 705/3; 789/1; 864/3; 940/3; 974/3; 1064/2;

Alignment Scores:

Pred. No.:	8.21	Length:	2329
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-19 (1-714) x T28125 (1-2329)

OY 412 AGATTTTAATCTCAATTTTA 435
Db 1632 ArgPheLeuIleSerLeuIleLeu 1639

RESULT 10

S34538

hypothetical protein 65 (rps18 3' region) - Euglena gracilis chloroplast

C:Species: chloroplast Euglena gracilis

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 29-Sep-1999

C:Accession: S34538; S34905

R:Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.; Monfort, A.; Orsat, B.; Spielmann,

submitted to the EMBL Data Library, January 1993

A:Description: The complete sequence of the Euglena gracilis chloroplast genome (tentati

A:Reference number: S34494

A:Accession: S34538

A:Molecule type: DNA

A:Residues: 1-65 <HAL1>

A:Cross-references: EMBL:X70810; NID:g415327; PIDN:CAA50118.1; PID:g415774

R:Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.R.; Monfort, A.; Orsat, B.; Spielman

Nucleic Acids Res. 21, 3537-3544, 1993

A:Title: Complete sequence of Euglena gracilis chloroplast DNA.

A:Reference number: S34862; MUID:93347989; PMID:8346031

A:Accession: S34905

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-65 <HAL2>

A:Cross-references: EMBL:X70810; NID:g415327; PIDN:CAA50118.1; PID:g415774

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

C:Genetics:

A:Genome: chloroplast

C:Superfamily: conserved hypothetical protein ycf9

C:Keywords: chloroplast

Alignment Scores:

Pred. No.:	140	Length:	65
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.23%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-19 (1-714) x S34538 (1-65)

OY 114 TTTTCCTAGTTTCTAATA 94
Db 51 PhePheLeuValPheLeuIle 57

RESULT 11

S21595

I9 heavy chain V region (7F2) - mouse

C:Species: Mus musculus (house mouse)

C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999

C:Accession: S21595

R:Kaartinen, M.

submitted to the EMBL Data Library, May 1992

A:Reference number: S21591

A:Accession: S21595

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-86 <KAA>

A:Cross-references: EMBL:X66458; NID:g51722; PIDN:CAA47073.1; PID:g51723

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Alignment Scores:

Pred. No.:	136	Length:	86
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.23%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-19 (1-714) x S21595 (1-86)

OY 680 TATGTAATATGTCGGG 660
Db 76 TyrGlyAsnTyrGlyLeuGly 82

RESULT 12

E84075

hypothetical protein BH3405 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: E84075

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: E84075

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-105 <STO>

A:Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BAB07124.1; GSPDB:G

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH3405

Alignment Scores:
Pred. No.: 133 Length: 105
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-19 (1-714) x E84075 (1-105)

OY 624 TTACTGTTCCCTAATTATTC 644
|||||
Db 11 LeuLeuPheLeuLeuPhe 17

RESULT 13

hypothetical protein aq_465 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: C70342
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: C70342
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-113 <AQF>
A:Cross-references: GB:AE000691; NID:g2983118; PIDN:AAC06729.1; PID:g2983125; GB:AE00065
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_465
C:Superfamily: Aquifex aeolicus hypothetical protein aq_465

Alignment Scores:

Pred. No.: 132 Length: 113
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-19 (1-714) x C70342 (1-113)

OY 428 AGTGAGATTAAATCTAAA 408
|||||
Db 59 SerGluIleLysAsnLeuLys 65

RESULT 14

T30769
hypothetical protein 8L - vaccinia virus (strain Ankara)
N:Alternate names: putative 13.7k protein
C:Species: vaccinia virus
A:Variety: strain Ankara
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T30769
R:Antoine, G.; Scheifflinger, F.; Falkner, F.G.; Dorner, F.
submitted to the EMBL Data Library, March 1997
A:Description: The complete genomic sequence of the modified Vaccinia Ankara (MVA) strain
A:Reference number: z20877
A:Accession: T30769
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-120 <ANT>
A:Cross-references: EMBL:U94848; PIDN:AAB96403.1
A:Experimental source: strain Ankara
C:Genetics:
A>Note: MVA008L

Alignment Scores:
Pred. No.: 131 Length: 120

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-19 (1-714) x T30769 (1-120)

OY 374 AAATACTGTTCTGATAGCA 394
|||||
Db 2 LysIleLeuPheLeuIleAla 8

RESULT 15

F72150
B6L protein - variola minor virus (strain Garcia-1966)
C:Species: variola minor virus
C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
C:Accession: F72150
R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Sifronov, P.F.; Massung, R.F.; Lo
submitted to GenBank, March 1998
A:Description: Analysis of the complete coding sequence of DNA of alastrim variola ml
A:Reference number: A72150
A:Accession: F72150
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-126 <SHC>
A:Cross-references: GB:Y16780; NID:g5830555; PIDN:CAB54600.1; PID:e1542556; PID:g5830
A:Experimental source: strain Garcia-1966
C:Genetics:
A:Gene: B6L

Alignment Scores:

Pred. No.: 130 Length: 126
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-19 (1-714) x F72150 (1-126)

OY 374 AAATACTGTTCTGATAGCA 394
|||||
Db 2 LysIleLeuPheLeuIleAla 8

Search completed: January 14, 2003, 17:40:35
Job time : 29.925 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:54 ; Search time 11.0253 Seconds
(without alignments)
5372.014 Million cell updates/sec

Title: US-09-835-992A-19

Perfect score: 218
Sequence: 1 cgccagaaaaagtatttta.....cacagttgttgccaattca 714

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:
-MODEL-frame+_n2p.model -DEV-xlp
-Q-/cgn2_1/USPTO.epool/US09835992/runat_14012003_161525_4014/app_query.fasta_1.3932
-DB-SwissProt_40 -QFMT-fastan -SUFFIX-olig.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.cdl
-LIST=45 -DOCALLIGN=200 -THR_SCORE-quality -THR_MIN=1 -ALIGN=15 -MODE-LOCAL
-OUTFMT=ptc -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09835992.qcgn_1_1_77_@runat_14012003_161525_4014 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.7	208	1	PAFR_MACMU
2	8	3.7	342	1	PAFR_HUMAN
3	2	3.7	393	1	Y306_MYCGE
4	8	3.7	3178	1	YS89_CAEEL
5	5	3.2	65	1	YCF9_EUGGR
6	7	3.2	113	1	Y465_AQUAE
7	7	3.2	164	1	2S51_ARATH
8	7	3.2	164	1	2S53_ARATH
9	7	3.2	180	1	YPKW_THECU
10	7	3.2	194	1	YPBH_BAGSU
11	7	3.2	201	1	YAHB_ECOLI
12	7	3.2	214	1	Y540_AQUAE
13	7	3.2	245	1	FCEI_RAT
14	7	3.2	267	1	PYRF_PICST
15	7	3.2	268	1	PYRF_CANTR
16	7	3.2	278	1	HBD_DEIRA
17	7	3.2	278	1	Y110_MYCGE
18	7	3.2	307	1	NADA_METJA

C	19	7	3.2	314	1	RLUC_BUCAL	P57430 buchnera ap
	20	7	3.2	314	1	Y315_MYCPN	Q50362 mycoplasma
	21	7	3.2	326	1	VS09_ROT2	Q03874 equine rota
	22	7	3.2	326	1	VS09_ROT2R	P11855 human rotav
	23	7	3.2	326	1	VS09_ROT2H	P12476 rhesus rota
	24	7	3.2	328	1	VMSA_HPBUD	P03145 duck hepati
	25	7	3.2	345	1	CLT2_PIG	Q95n03 sus scrofa
	26	7	3.2	365	1	VMSA_HPBDC	P30029 duck hepati
	27	7	3.2	366	1	VMSA_HPBDB	P17194 duck hepati
	28	7	3.2	366	1	VMSA_HPBWD	P17195 duck hepati
	29	7	3.2	379	1	PANE_YEAST	P38787 saccharomyc
	30	7	3.2	380	1	CAPM_STAUA	P39862 staphylococ
	31	7	3.2	387	1	YG26_YEAST	P53063 saccharomyc
C	32	7	3.2	397	1	PGK_AQUAE	O66519 aquifex aeo
	33	7	3.2	405	1	DXR_PASMU	P57985 pasteurella
	34	7	3.2	423	1	MTB5_NEIGO	Q59605 neisseria g
	35	7	3.2	423	1	MTN4_NEILA	P50182 neisseria l
C	36	7	3.2	430	1	DIN7_YEAST	Q12086 saccharomyc
	37	7	3.2	433	1	ENO_VIBCH	Q9kpc5 vibrio chol
C	38	7	3.2	436	1	HMDH_ARCFU	O28538 archaeoglob
	39	7	3.2	460	1	YAS4_HAEIN	P44104 haemophilus
	40	7	3.2	473	1	IAIC_MALDO	P37821 malus domes
	41	7	3.2	527	1	Y099_CAEEL	P41849 caenorhabdl
C	42	7	3.2	560	1	YJK5_YEAST	P42948 saccharomyc
	43	7	3.2	580	1	RGPI_XENLA	O13066 xenopus lae
	44	7	3.2	678	1	GYRB_MYCLE	Q59533 mycobacteri
	45	7	3.2	681	1	VGP_MABVM	P35253 marburg vir

ALIGNMENTS

RESULT 1
PAFR_MACMU
ID PAFR_MACMU STANDARD; PRT; 208 AA.
AC P35366;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Platelet activating factor receptor (PAF-R) (Fragment).
GN PTAFR.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Behal R.H., Debussere M.S., Olson M.S.;
RL Submitted (xxx-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR PLATELET ACTIVATING FACTOR, A CHEMOTACTIC
CC PHOSPHOLIPID MEDIATOR THAT POSSESSES POTENT INFLAMMATORY, SMOOTH-
CC MUSCLE CONTRACTILE AND HYPOTENSIVE ACTIVITY. SEEM TO MEDIATE ITS
CC ACTION VIA A G PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L07333; AAA68893.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Chemotaxis.
FT DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).

```
FT TRANSMEM 17 38 1 (POTENTIAL).
FT DOMAIN 39 54 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 55 74 2 (POTENTIAL).
FT DOMAIN 75 91 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 92 113 3 (POTENTIAL).
FT DOMAIN 114 133 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 134 155 4 (POTENTIAL).
FT DOMAIN 156 184 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 185 205 5 (POTENTIAL).
FT DOMAIN 206 >208 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 90 173 BY SIMILARITY.
FT NON_TER 208 208
SQ SEQUENCE 208 AA; 23736 MW; 7985928F70B3C6A1 CRC64;

Alignment Scores:
Pred. No.: 11 Length: 208
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-19 (1-714) x PAFR_MACMU (1-208)

QY 117 TCTTTTTCCTAGTTTCTAATA 94
Db 194 SerpPheLeuValPheLeuIle 201

RESULT 2
PAFR_HUMAN
ID PAFR_HUMAN STANDARD; PRT; 342 AA.
AC P25105;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Platelet activating factor receptor (PAF-R).
GN PTAFR OR PAFR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92028922; Pubmed=1656963;
RA Ye R.D., Prossnitz E.R., Zou A., Cochran C.G.;
RT "Characterization of a human cDNA that encodes a functional receptor
RL for platelet activating factor."
RN Biochem. Biophys. Res. Commun. 180:105-111(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
RX MEDLINE=92041873; Pubmed=1657923;
RA Nakamura M., Honda Z., Izumi T., Sakanaka C., Mutoh H., Minami M.,
RT Bito H., Seyama Y., Matsumoto T., Noma M., Shimizu T.;
RT "Molecular cloning and expression of platelet-activating factor
RL receptor from human leukocytes."
RN J. Biol. Chem. 266:20400-20405(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92250505; Pubmed=1374385;
RA Kunz D., Gerard N.P., Gerard C.;
RT "The human leukocyte platelet-activating factor receptor. cDNA
RT cloning, cell surface expression, and construction of a novel
RT epitope-bearing analog."
RL J. Biol. Chem. 267:9101-9106(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92347886; Pubmed=1322356;
RA Seyfried C.E., Schweickart V.L., Godiska R., Gray P.W.;
RT "The human platelet-activating factor receptor gene (PTAFR) contains
RT no introns and maps to chromosome 1."
RL Genomics 13:832-834(1992).
```

```
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=93112021; Pubmed=1281995;
RA Sugimoto T., Tsuchimochi H., McGregor C.G., Mutoh H., Shimizu T.,
RA Kurachi Y.;
RT "Molecular cloning and characterization of the platelet-activating
RT factor receptor gene expressed in the human heart."
RL Biochem. Biophys. Res. Commun. 189:617-624(1992).
RN [6]
RP SEQUENCE FROM N.A.
RA Behal R.H., Debussere M.S., Olson M.S.;
RL Submitted (XXX-1992) to the EMBL/Genbank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=93192035; Pubmed=8383507;
RA Chase P.B., Halonen M., Regan J.W.;
RT "Cloning of a human platelet-activating factor receptor gene:
RT evidence for an intron in the 5'-untranslated region."
RL Am. J. Respir. Cell Mol. Biol. 8:240-244(1993).
CC -1- FUNCTION: RECEPTOR FOR PLATELET ACTIVATING FACTOR, A CHEMOTACTIC
CC PHOSPHOLIPID MEDIATOR THAT POSSESSES POTENT INFLAMMATORY, SMOOTH-
CC MUSCLE CONTRACTILE AND HYPOTENSIVE ACTIVITY. SEEM TO MEDIATE ITS
CC ACTION VIA A G PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M80436; AAA60001.1; -
DR EMBL; M76674; AAA60002.1; -
DR EMBL; D10202; BA01050.1; -
DR EMBL; M88177; AAA60214.1; -
DR EMBL; S52624; AAB24695.2; -
DR EMBL; I07334; AAA60108.1; -
DR EMBL; S56396; AAB25755.1; -
DR PIR; JH0479; JH0479.
DR PIR; A40191; A40191.
DR PIR; A41079; A41079.
DR Gene; HGNC:9582; PTAFR.
DR MIM; 173393; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECIP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECIP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Chemotaxis;
KW Polymorphism.
FT DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 17 38 1 (POTENTIAL).
FT DOMAIN 39 54 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 55 74 2 (POTENTIAL).
FT DOMAIN 75 91 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 92 113 3 (POTENTIAL).
FT DOMAIN 114 133 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 134 155 4 (POTENTIAL).
FT DOMAIN 156 184 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 185 205 5 (POTENTIAL).
FT DOMAIN 206 233 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 234 254 6 (POTENTIAL).
FT DOMAIN 255 276 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 277 296 7 (POTENTIAL).
FT DOMAIN 297 342 CYTOPLASMIC (POTENTIAL).
FT DISULFID 90 173 BY SIMILARITY.
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 224 224 A -> D (IN DBSNP:5938).
FT /FTID=VAR_011851.
```

FT VARIANT 338 338 N -> S (IN DBSNP:5939).
FT CONFLICT 28 28 /FTID-VAR_011852.
FT CONFLICT 66 66 L -> P (IN REF. 6).
FT CONFLICT 95 95 F -> L (IN REF. 6).
FT CONFLICT 227 228 C -> R (IN REF. 6).
FT CONFLICT 227 228 KR -> TG (IN REF. 4).
FT CONFLICT 247 228 KR -> TT (IN REF. 6).
FT CONFLICT 316 247 P -> A (IN REF. 6).
FT CONFLICT 316 316 K -> N (IN REF. 5).
SQ SEQUENCE 342 AA; 39203 MW; 890073C9EBA79228 CRC64;

Alignment Scores:

Pred. No.: 9.88 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-19 (1-714) x PAFR_HUMAN (1-342)

OY 117 TCTTTTCTAGTTTCTAATA 94
Db 194 SerpPheLeuValPheLeuIle 201

RESULT 3

Y306_MYCGE STANDARD; PRT; 393 AA.
ID Y306_MYCGE
AC P47548;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG306.
GN MG306.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID-2097;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE-96026346; PubMed-7569993;
RA Fraser C.M., Cocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudak D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; U39711; AAC71528.1; -
DR TIGR; MG306; -
KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 131 151 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 253 273 POTENTIAL.
FT TRANSMEM 289 309 POTENTIAL.
FT TRANSMEM 349 369 POTENTIAL.
SQ SEQUENCE 393 AA; 45750 MW; D740FDA979EC364A CRC64;

Alignment Scores:

Pred. No.: 9.6 Length: 393
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-19 (1-714) x Y306_MYCGE (1-393)

OY 117 TCTTTTCTAGTTTCTAATA 94
Db 19 SerpPheLeuValPheLeuIle 26

RESULT 4

YS89_CAEEL STANDARD; PRT; 3178 AA.
ID YS89_CAEEL
AC Q09624; Q09625; Q969D4;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ZK945.9 in chromosome II.
GN ZK945.9/ZK945.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID-6239;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Wilkinson-Spratt J.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [2]

RP REVISIONS.
RA Durbin R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; Z48544; CAB70192.1; -
DR EMBL; Z48582; CAB70192.1; JOINED.
DR EMBL; Z48582; CAB70201.1; -
DR EMBL; Z48544; CAB70201.1; JOINED.
DR WormPep; ZK945.9; CE25697.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR001024; Lipoxxygenase_LH2.
DR InterPro; IPR000636; M+channel_nlg.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF01477; PLAT; 1.
DR Pfam; PF01825; GPS; 1.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00308; LH2; 1.

DR Hypothetical protein; Transmembrane.
KW Hypothetical protein; Transmembrane.
FT DOMAIN 266 1196 SER/THR-RICH.
FT DOMAIN 1105 1241 GLY/SER-RICH.
FT DOMAIN 2071 2120 GPS.
FT TRANSMEM 13 30 POTENTIAL.
FT TRANSMEM 51 73 POTENTIAL.
FT TRANSMEM 2139 2161 POTENTIAL.
FT TRANSMEM 2348 2367 POTENTIAL.
FT TRANSMEM 2390 2412 POTENTIAL.
FT TRANSMEM 2451 2468 POTENTIAL.
FT TRANSMEM 2483 2505 POTENTIAL.
FT TRANSMEM 2567 2589 POTENTIAL.
FT TRANSMEM 2836 2858 POTENTIAL.

FT TRANSMEM 2939 2961 POTENTIAL.
FT TRANSMEM 2976 2998 POTENTIAL.
FT TRANSMEM 3038 3060 POTENTIAL.
SQ SEQUENCE 3178 AA; 344726 MW; F8239436D03666CD CRC64;

Alignment Scores:

Pred. No.: 6 17 Length: 3178
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-19 (1-714) x YS89_CAEEL (1-3178)

OY 412 AGATTTTAACTCAGCTAATTTTA 435

Db 2481 ArgPheLeuIleSerLeuIleLeu 2488

RESULT 5

YCF9_EUGGR STANDARD; PRT; 65 AA.
ID YCF9_EUGGR

AC P32095;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 7.6 kDa protein ycf9 (ORF 65).

GN YCF9.

OS Euglena gracilis.

OG Chloroplast.

OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.

OX NCBI_TaxID=3039;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Z;

RX MEDLINE=93347989; PubMed=8346031;

RA Hallick R.B., Hong L., Drager R.G., Favreau M.R., Monfort A.,

RA Orsat B., Spielmann A., Stutz E.;

RT "Complete sequence of Euglena gracilis chloroplast DNA.";

RL Nucleic Acids Res. 21:3537-3544(1993).

CC -1- SIMILARITY: BELONGS TO THE YCF9 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; 211874; -; NOT_ANNOTATED_CDS.

DR EMBL; X70810; CA50118.1; -.

DR PIR; S34538; S34538.

DR InterPro; IPR002644; ycf9_struct.

DR Pfam; PF01737; YCF9; 1.

DR ProDom; PD004770; ycf9_struct; 1.

KW Chloroplast; Hypothetical protein.

SQ SEQUENCE 65 AA; 7586 MW; 67DE2359D7A0F771 CRC64;

Alignment Scores:

Pred. No.: 138 Length: 65
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-19 (1-714) x YCF9_EUGGR (1-65)

OY 114 TTTTCTAGTTTCTCTATA 94

Db 51 PhePheLeuValPheLeuIle 57

RESULT 6

Y465_AQUAE STANDARD; PRT; 113 AA.
ID Y465_AQUAE

AC O66767;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein AQ_465.

GN AQ_465.

OS Aquifex aeolicus.

OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;

OC Aquifex.

OX NCBI_TaxID=63363;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VF5;

RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,

RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,

RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex

RT aeolicus.";

RL Nature 392:353-358(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE000691; AAC06729.1; -.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 7 29 POTENTIAL.

SQ SEQUENCE 113 AA; 13252 MW; F402B9193B2A8522 CRC64;

Alignment Scores:

Pred. No.: 123 Length: 113
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-19 (1-714) x Y465_AQUAE (1-113)

OY 428 AGTGAGTTAAATCTAATA 408

Db 59 SerGluIleLysAsnLeuLys 65

RESULT 7

2S51_ARATH STANDARD; PRT; 164 AA.
ID 2S51_ARATH

AC P15457;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 2S seed storage protein 1 precursor (2S albumin storage protein)

DE (NMWU2-2S albumin 1).

GN AT2S1 OR AT4G27140 OR T24A18.90.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 38-73 AND 84-162.

RC STRAIN=cv. C24;

RA Krebbers E., Herdies L., de Clercq A., Seurinck J., Iemans J.,

RA van Damme J., Segura M., Gheysen G., van Montagu M.,

RA Vandekerckhove J.;

RT "Determination of the processing sites of an Arabidopsis 2S albumin

RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohnselt J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weltzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berneriser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
RA Pettelt A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argitieu A., Vitale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Sheker M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Maira M., Martensen R., McCombie W.R.,
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana."
RL Nature 402:769-777(1999).
RN [4]
RP SEQUENCE OF 103-164 FROM N.A.
RC STRAIN=cv. Columbia;
RA Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.,
RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
CC LINKED BY DISULFIDE BONDS.
CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22035; AAA32745.1; -;
DR EMBL; 224744; CA80868.1; -;
DR EMBL; AL035680; CAB38846.1; -;
DR EMBL; AL161566; CAB79571.1; -;
DR EMBL; Z17580; CAA79001.1; -;
DR PIR; JAO163; NMMU3.
DR PIR; S34674; S34674.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR000617; Napin.
DR InterPro; IPR001768; Tryp/amy1_inhbt.
DR Pfam; PF00234; tryp_alpha_amy1; 1.
DR PRINTS; PR00496; NAPIN.
DR ProDom; PD002498; Napin; 1.

DR SMART; SM00499; AAI; 1.
KW Seed storage protein; Albumin; Signal; Multigene family.
FT SIGNAL 1 21
FT PROPEP 22 37
FT CHAIN 38 72 2S SEED STORAGE PROTEIN 3 SMALL SUBUNIT
FT PROPEP 73 81 (BY SIMILARITY).
FT CHAIN 82 164
FT SEQUENCE 164 AA; 18762 MW; 7399D55F9E204D40 CRC64;
SQ (BY SIMILARITY).
Alignment Scores:
Pred. No.: 114 Length: 164
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 1 Gaps: 0
US-09-835-992A-19 (1-714) x 2SS3_ARATH (1-164)
QY 496 TTGGCATGTGTCTTACTT 516
Db 12 LeuAlaLeuCysPheLeuLeu 18
RESULT 9
YPKW_THECU
ID YPKW_THECU STANDARD; PRT; 180 AA.
AC P49694;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein in pkwa 5'region (ORF1) (Fragment).
OS Thermomonospora curvata.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinobacteriales;
OC Actinomycetales; Streptosporangineae; Thermomonosporaceae;
OC Thermomonospora.
OX NCBI_TaxID=2020;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM 3352;
RX MEDLINE=96200125; PubMed=8631732;
RA Janda L., Tichy P., Spizek J., Petricek M.,
RT "A deduced Thermomonospora curvata protein containing
RT serine/threonine protein kinase and WD-repeat domains.";
RL J. Bacteriol. 178:1487-1489(1996).
CC -1- SIMILARITY: BELONGS TO THE CDAR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF115313; -; NOT_ANNOTATED_CDS.
DR InterPro; IPR000712; Bcl2_BH.
KW Hypothetical protein.
FT NON_TER 1
FT SEQUENCE 180 AA; 19880 MW; 7399D55F9E204D40 CRC64;
SQ
Alignment Scores:
Pred. No.: 112 Length: 180
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 1 Gaps: 0
US-09-835-992A-19 (1-714) x YPKW_THECU (1-180)
QY 76 AATTCAGTGAGATATAGATA 56

Db 132 AsnServAlargTyrArgIle 138

RESULT 10

YPBH_BACSU STANDARD; PRT; 194 AA.

ID YPBH_BACSU

AC P50734;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein ypbh.

CN YPBH.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-168 / Marburg;

RX MEDLINE=96349105; PubMed=8760912;

RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,

RA Serror P.;

RT "Sequence analysis of the Bacillus subtilis chromosome region between

RT the *serA* and *kds* loci cloned in a yeast artificial chromosome.";

RL Microbiology 142:2005-2016(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

RA Entian K.D., Erlington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koeltter P., Konlingstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadate Y.,

RA Sato T., Scanlan E., Schlecht S., Schroeter R., Scoffone F.,

RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambuit R., Wedler E., Wedler H., Weltzenegger T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

RT "The complete genome sequence of the Gram-positive bacterium Bacillus

RT subtilis.";

RL Nature 390:249-256(1997).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; L47648; AAC83952.1; -

DR EMBL; 299115; CAB14213.1; -

DR EMBL; 299116; CAB14229.1; -

DR Subtilist; BG11434; ypbh.

KW Hypothetical protein; Complete proteome.

SO SEQUENCE 194 AA; 22157 MW; B27C6ABE4EA0745C CRC64;

Alignment Scores:

Pred. No.: 110

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 3.23%

DB: 1

Gaps: 0

US-09-835-992a-19 (1-714) x YPBH_BACSU (1-194)

OY 317 AAAATATAATATTTTAAACA 297

Db 10 LysileLysilePheLeuThr 16

RESULT 11

YAHB_ECOLI

ID YAHB_ECOLI STANDARD; PRT; 201 AA.

AC P77736;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical ANK-repeats protein yahb.

CN YAHB OR B0318.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [2]

RP SEQUENCE FROM N.A.

RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,

RA Federspiel N., Hyman R., Kallman R., Komp C., Kurdi O., Lew H., Lin D.,

RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AE000139; AAC73421.1; -

DR EMBL; U73857; AAB18044.1; -

DR Ecogene; EG13588; yahb.

DR InterPro; IPR002110; ANK.

DR Pfam; PF00023; ank; 6.

DR SMART; SM00248; ANK; 2.

DR PROSITE; PSS0088; ANK_REPEAT; 3.

DR PROSITE; PSS0297; ANK_REPEAT_REGION; 1.

KW Hypothetical protein; Repeat; ANK repeat; Complete proteome.

FT REPEAT 5 34 ANK 1.

FT REPEAT 38 67 ANK 2.

FT REPEAT 71 100 ANK 3.

FT REPEAT 104 134 ANK 4.

FT REPEAT 138 172 ANK 5.

FT REPEAT 176 201 ANK 6.

SO SEQUENCE 201 AA; 21687 MW; A90244CEB892E8BD CRC64;

Alignment Scores:

Pred. No.: 109 Length: 201
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-19 (1-714) x YAHU_ECOLI (1-201)

QY 326 GCCTGTGTCAAGCTGATC 346

DB 53 AlacysvalGlnAlaLeuile 59

RESULT 12

Y540_AOUAE STANDARD; PRT; 214 AA.
ID Y540_AOUAE

AC 066819;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein AQ_540.

GN AQ_540.

OS Aquifex aeolicus.

OC Bacteria; Aquificae; Aquificales; Aquificaceae;

OC Aquifex.

OX NCBI_TaxID=63363;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VF5;

RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,

RA Graham D.E., Overbeek R., Snead M.A., Keller M., Anjay M., Huber R.,

RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex

RT aeolicus.";

RL Nature 392:353-358(1998).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: BELONGS TO THE UPF0056 (MARC) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE000694; AAC06780.1; -.

DR InterPro: IPR002771; MARC.

DR Pfam: PF01914; MARC; 1.

DR TIGRFAMs: TIGR00427; UPF0056; 1.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 17

FT TRANSMEM 47

FT TRANSMEM 73

FT TRANSMEM 122

FT TRANSMEM 153

FT TRANSMEM 185

FT TRANSMEM 205

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

FT TRANSMEM 214

RESULT 13

FCE1_RAT STANDARD; PRT; 245 AA.
ID FCE1_RAT

AC P12371;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE High affinity immunoglobulin epsilon receptor alpha-subunit precursor

DE (FCER1) (IGE FC receptor, alpha-subunit) (FC-epsilon RI-alpha).

GN FCER1A OR FCE1A.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88024987; PubMed=2959318;

RA Kinet J.-P., Metzger H., Hakimi J., Kochan J.;

RT "A CDNA presumably coding for the alpha subunit of the receptor

RT with high affinity for immunoglobulin E.";

RL Biochemistry 26:4605-4610(1987).

RN [2]

RP REVISIONS.

RA Kochan J.;

RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Mast cells;

RX MEDLINE=88158102; PubMed=2964640;

RA Shimizu A., Tepler I., Benfey P.N., Berenstein E.H., Siraganian R.P.,

RA Leder P.;

RT "Human and rat mast cell high-affinity immunoglobulin E receptors:

RT characterization of putative alpha-chain gene products.";

RL Proc. Natl. Acad. Sci. U.S.A. 85:1907-1911(1988).

RN [4]

RP SEQUENCE OF 21-245 FROM N.A.

RX MEDLINE=88289772; PubMed=2969594;

RA Liu F.-T., Albrandt K., Robertson M.W.;

RT "CDNA heterogeneity suggests structural variants related to the high-

RT affinity IGE receptor.";

RL Proc. Natl. Acad. Sci. U.S.A. 85:5639-5643(1988).

CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH

CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC

CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL

CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)

CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR

CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.

CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO

CC DISULFIDE LINKED GAMMA CHAINS.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M17153; AAAA2045.1; -.

DR EMBL: J03606; AAAA41582.1; -.

DR EMBL: M21622; AAAA1146.1; -.

DR PIR: A27116; A27116.

DR PIR: A30154; A30154.

DR PIR: C31327; C31327.

DR HSSP: P12319; 1ALS.

DR InterPro: IPR003599; Ig.

DR InterPro: IPR003006; Ig_MHC.

DR Pfam: PF00047; Ig; 2.

DR SMART: SM00409; Ig; 2.

KW IGE-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;

KW Immunoglobulin domain; Repeat.
FT SIGNAL 1 23
FT CHAIN 24 245
FT TRANSMEM 24 204
FT DOMAIN 205 223
FT DOMAIN 224 245
FT DOMAIN 42 98
FT DOMAIN 123 181
FT DISULFID 49 91
FT DISULFID 130 174
FT CARBOHYD 52 52
FT CARBOHYD 53 53
FT CARBOHYD 58 58
FT CARBOHYD 65 65
FT CARBOHYD 123 123
FT CARBOHYD 158 158
FT CARBOHYD 167 167
SQ SEQUENCE 245 AA; 27793 MW; A0E67DD363B72197 CRC64;

HIGH AFFINITY IMMUNOGLOBULIN EPSILON
RECEPTOR ALPHA-SUBUNIT.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
IG-LIKE DOMAIN 1.
IG-LIKE DOMAIN 2.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).

Alignment Scores:
Pred. No.: 105 Length: 245
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-19 (1-714) x FCEL_RAT (1-245)

OY 83 AATAACATTAGTAGAATAA 103
DB 157 AsnAsnIleSerIleArgLys 163

RESULT 14
PYRF_PICST STANDARD; PRT; 267 AA.
AC P49434;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase).
DE (OMPDCase) (OMPdecase) (Uridine 5'-monophosphate synthase) (UMP synthase).
DE URA3.
GN URA3.
OS Pichia stipitis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4924;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 58785 / CBS 6054;
RX MEDLINE=95110115; Pubmed=7811063;
RA Yang V.W., Marks J.A., Davis B.P., Jeffries T.W.;
RT "High-efficiency transformation of Pichia stipitis based on its URA3 gene and a homologous autonomous replication sequence, ARS2.";
RL Appl. Environ. Microbiol. 60:4245-4254(1994).
CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate -> UMP + CO(2).
CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U08629; AAA65978.1; -.
DR HSSP; P03962; IDOW.
DR InterPro; IPR001754; OMPdecase.
DR Pfam; PF00215; OMPdecase; 1.

DR PROSITE; PS00156; OMPDECASE; 1.
KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
FT ACT_SITE 94 94
SQ SEQUENCE 267 AA; 29529 MW; D98361C398F49E06 CRC64;

Alignment Scores:
Pred. No.: 103 Length: 267
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-19 (1-714) x PYRF_PICST (1-267)

OY 521 GAACCTAGTAAGAACACAAAT 501
DB 80 GluLeuSerLysLysHisAsn 86

RESULT 15

PYRF_CANTR STANDARD; PRT; 268 AA.
AC 042771;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
DE (OMPDCase) (OMPdecase) (Uridine 5'-monophosphate synthase) (UMP synthase).
DE URA3.
GN Candida tropicalis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mtosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5482;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-M4;
RX MEDLINE=98403409; Pubmed=9734982;
RA Su J.-H., Hsia J.-H., Chang M.-C.;
RT "Cloning and sequence analysis of the Candida tropicalis URA3 gene encoding orotidine-5'-phosphate decarboxylase.";
RL Curr. Microbiol. 37:210-213(1998).
CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate -> UMP + CO(2).
CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF040702; AAB96773.1; -.
DR HSSP; P03962; IDOW.
DR InterPro; IPR001754; OMPdecase.
DR Pfam; PF00215; OMPdecase; 1.
DR PROSITE; PS00156; OMPDECASE; 1.
KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
FT ACT_SITE 94 94
SQ SEQUENCE 268 AA; 29672 MW; A43C1ED1CB24237 CRC64;

Alignment Scores:
Pred. No.: 103 Length: 268
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-19 (1-714) x PYRF_CANTR (1-268)

OY 521 GAAGTAACTAAGAAACACAAT 501
|||||
Db 80 GluLeuSerLysLysHisAsn 86

Search completed: January 14, 2003, 17:29:07
Job time : 17.0253 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:20:59 ; Search time 50.2755 Seconds
(without alignments)
5852.461 Million cell updates/sec

Title: US-09-835-992A-19
Perfect score: 218
Sequence: 1 cgcacgaaaaagttatttta.....cacagttgttgccaattca 714

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1
Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09835992/runat_14012003_161526_4027/app_query.fasta_1.3932
-DB=SPTREMBL_21 -QFMT=fastan -SUFFIX=olig.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdt
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09835992_@CGN_1_1_346_@runat_14012003_161526_4027 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result				SUMMARIES	
No.	Score	Match	Length	ID	Description
c 1	8	3.7	123	17	Q973P2 sulfolobus

c	2	8	3.7	123	17	Q96Zw9	Q96Zw9 sulfolobus
c	3	8	3.7	151	6	Q97662	Q97662 ovis aries
c	4	8	3.7	195	17	O59004	O59004 pyrococcus
c	5	8	3.7	280	5	Q9N5G4	Q9N5G4 caenorhabdl
c	6	8	3.7	342	6	Q9GK76	Q9GK76 capra hircu
c	7	8	3.7	342	6	Q9TTY5	Q9TTY5 bos taurus
c	8	8	3.7	343	5	O17820	Q17820 caenorhabdl
c	9	8	3.7	352	17	O8TTZ5	Q8ttz5 methanosarc
c	10	8	3.7	375	16	Q8XN39	Q8xn39 clostridium
c	11	8	3.7	376	16	Q92RV3	Q92rv3 rhizobium m
c	12	8	3.7	413	17	Q972W9	Q972w9 sulfolobus
c	13	8	3.7	485	11	Q9JM15	Q9jml5 ratius norv
c	14	8	3.7	485	11	Q99PR1	Q99pr1 mus musculu
c	15	8	3.7	512	16	Q9HVA7	Q9hva7 pseudomonas
c	16	8	3.7	791	17	Q9YB89	Q9yb89 aeropyrum p
c	17	8	3.7	1115	4	Q9HD43	Q9hd43 homo sapien
c	18	8	3.7	1118	4	O15426	Q15426 homo sapien
c	19	8	3.7	1183	16	Q98QH9	Q98qh9 mycoplasma
c	20	8	3.7	3178	5	Q969D4	Q969d4 caenorhabdl
c	21	7	3.2	55	10	Q41910	Q41910 arabidopsis
c	22	7	3.2	55	12	Q91FK1	Q91fk1 chilo lride
c	23	7	3.2	72	6	Q9GMK4	Q9gmk4 macaca fasc
c	24	7	3.2	85	5	Q9VDM5	Q9vdm5 drosophila
c	25	7	3.2	105	16	Q9K7F6	Q9k7f6 bacillus ha
c	26	7	3.2	106	2	Q47966	Q47966 herpetosiph
c	27	7	3.2	117	8	O09313	O09313 nesticus st
c	28	7	3.2	117	8	O09314	O09314 nesticus st
c	29	7	3.2	117	8	O09317	O09317 nesticus ba
c	30	7	3.2	117	8	O09333	O09333 nesticus sl
c	31	7	3.2	120	12	O57169	O57169 vaccinia vl
c	32	7	3.2	121	2	Q9AL17	Q9al17 enterococcu
c	33	7	3.2	124	12	P87608	P87608 cowpox viru
c	34	7	3.2	126	12	O89067	O89067 variola vir
c	35	7	3.2	139	8	Q951Z7	Q951z7 neoneurus m
c	36	7	3.2	167	12	Q9WFA4	Q9wfa4 snow goose
c	37	7	3.2	167	12	Q9WEA7	Q9wea7 snow goose
c	38	7	3.2	167	12	Q9WFB0	Q9wfb0 snow goose
c	39	7	3.2	167	12	Q9WFB4	Q9wfb4 snow goose
c	40	7	3.2	167	12	Q9WFB7	Q9wfb7 snow goose
c	41	7	3.2	167	12	O90742	O90742 duck hepati
c	42	7	3.2	167	12	O89748	O89748 duck hepati
c	43	7	3.2	167	12	O89689	O89689 duck hepati
c	44	7	3.2	181	8	Q954J1	Q954j1 panoplothr
c	45	7	3.2	183	10	Q8RWP7	Q8rwp7 arabidopsis

ALIGNMENTS

RESULT 1			
Q973P2	Q973P2	PRELIMINARY;	PRT; 123 AA.
AC	Q973P2;		
DT	01-DEC-2001 (TREMblrel. 19, Created)		
DT	01-DEC-2001 (TREMblrel. 19, Last sequence update)		
DT	01-DEC-2001 (TREMblrel. 19, Last annotation update)		
DE	Putative cryptogene protein G4.		
GN	ST0856.		
OS	Sulfolobus tokodall.		
OC	Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;		
OC	Sulfolobus.		
OX	NCBI_TaxID=111955;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-JCM 10545 / 7;		
RX	PubMed=11572479;		
RA	Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,		
RA	Sekine M., Baba S.-I., Ankal A., Kosugi H., Hosoyama A., Fukui S.,		
RA	Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,		
RA	Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,		
RA	Aoki K.-I., Masuda S., Yanagil M., Nishimura M., Yamagishi A.,		
RA	Oshima T., Kikuchi H.;		
RT	"Complete genome sequence of an aerobic thermophilic		
RT	Crenarchaeon, Sulfolobus tokodall strain7.";		

RL DNA Res. 8:123-140(2001).
DR EMBL; AP000984; BAB65869.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 123 AA; 14153 MW; 4C04001E122B0D02 CRC64;

Alignment Scores:

Pred. No.:	35.6	Length:	123
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	17	Gaps:	0

US-09-835-992A-19 (1-714) x Q973P2 (1-123)

QY 120 TTATCTTTTTCCTAGTTTCTA 97
|||||
Db 106 LeuSerPhepheLeuValPheLeu 113

RESULT 2

Q96ZW9 PRELIMINARY; PRT; 123 AA.

AC Q96ZW9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein ST1718.
GN ST1718.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000987; BAB66804.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 123 AA; 14210 MW; D4AD4AC9B8A5826F CRC64;

Alignment Scores:

Pred. No.:	35.6	Length:	123
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	17	Gaps:	0

US-09-835-992A-19 (1-714) x Q96ZW9 (1-123)

QY 120 TTATCTTTTTCCTAGTTTCTA 97
|||||
Db 106 LeuSerPhepheLeuValPheLeu 113

RESULT 3

Q97662 PRELIMINARY; PRT; 151 AA.

AC Q97662;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Platelet activating factor receptor (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;
RA Sander F.C., Ibe O.B., Raj J.U.;
RT "Developmentally Regulated Expression of Platelet Activating Factor
RT Receptor in Ovine Lung."
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF099674; AAC77459.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.

KW Receptor.
FT NON_TER 1 1
FT NON_TER 151 151
SQ SEQUENCE 151 AA; 17001 MW; 31EC4F2006AF951C CRC64;

Alignment Scores:

Pred. No.:	34.1	Length:	151
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	6	Gaps:	0

US-09-835-992A-19 (1-714) x Q97662 (1-151)

QY 117 TCTTTTTCCTAGTTTCTATA 94
|||||
Db 138 SerPhepheLeuValPheLeuIle 145

RESULT 4

O59004 PRELIMINARY; PRT; 195 AA.

AC O59004;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein PH1301.
GN PH1301.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL; AP000006; BAA30405.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 195 AA; 22883 MW; C1075554545837CC CRC64;

Alignment Scores:

Pred. No.:	32.2	Length:	195
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	17	Gaps:	0

US-09-835-992A-19 (1-714) x 059004 (1-195)

OY 120 TTATCTTTTCTAGTTTCTA 97

DB 99 LeuScrpheleuValpheleu 106

RESULT 5

O9N5G4 PRELIMINARY; PRT; 280 AA.

AC O9N5G4; 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 29.9 kDa protein.
GN R05D8.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Langston Y., Wohlmann P., Lennox S.;
RT "The sequence of C. elegans cosmid R05D8.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
(SDR) FAMILY.
CC EMBL: AC006677; AAF39950.1; -.
DR HSSP; P50162; IAE1.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Hypothetical protein; Oxidoreductase.
SQ SEQUENCE 280 AA; 29923 MW; 5B8C9AE6F1D5F006 CRC64;

Alignment Scores:

Pred. No.:	29.7	Length:	280
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	5	Gaps:	0

US-09-835-992A-19 (1-714) x O9N5G4 (1-280)

OY 531 AAAACTCTTAATGAACTTCAG 554

DB 116 LysThrLeuLysLeuAsnLeuGln 123

RESULT 6

O9GK76 PRELIMINARY; PRT; 342 AA.

AC O9GK76; 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Platelet-activating factor receptor.
OS Capra hircus (Goat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang W., Diehl J.R., Plum F.;
RT "Partial Sequence of Caprine Platelet-Activating Factor Receptor
Gene.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF302764; AAG39982.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR01559; DUFFYANTIGEN.
DR PROSITE; PS00237; GPCR_RHODPSN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 342 AA; 39612 MW; 39D9379AED60FDAE CRC64;

Alignment Scores:

Pred. No.:	28.5	Length:	342
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	6	Gaps:	0

US-09-835-992A-19 (1-714) x O9GK76 (1-342)

OY 117 TCCTTTTCTAGTTTCTAATA 94

DB 194 SerPhepheleuValpheleuile 201

RESULT 7

O9TTY5 PRELIMINARY; PRT; 342 AA.

AC O9TTY5; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Platelet-activating factor receptor.
GN PAFR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang W.S., Diehl J.R., Roudsbush W.E.;
RT "Partial Sequence of Bovine Platelet-Activating Factor Receptor
Gene.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA He B., Tiemann U., Kanitz W., Weikard R., Laurent P., Schwerin M.,
RA Schmidt P.;
RT "Molecular characterization of bovine platelet-activating factor
receptor transcripts and their detection in different tissues of
cattle.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF187321; AAF01439.2; -.
DR EMBL; AJ295321; CAC43290.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR01559; DUFFYANTIGEN.
DR PROSITE; PS00237; GPCR_RHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 342 AA; 39691 MW; 7C4236205AE937C9 CRC64;

Alignment Scores:

Pred. No.:	28.5	Length:	342
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	6	Gaps:	0

US-09-835-992A-19 (1-714) x Q9TTY5 (1-342)

QY 117 TCTTTTTCCTAGTTTCTAATA 94
|||||

Db 194 SerpHeLeuValPheLeuIle 201

RESULT 8

Q17820 PRELIMINARY; PRT; 343 AA.

AC Q17820; 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

GN C08B6.9 protein.

GN C08B6.9.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RL Wilkinson J.;

RN Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT "Genome sequence of the nematode C.elegans: A platform for

RT investigating biology.";

RL Science 282:2012-2018(1998).

DR EMBL; Z72502; CAA96591.1; -.

DR InterPro; IPR00594; ThIF_domain.

DR Pfam; PF00899; ThIF; 1.

SQ SEQUENCE 343 AA; 38738 MW; 2897C0F918E81B6 CRC64;

Alignment Scores:

Pred. No.:	28.4	Length:	343
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	5	Gaps:	0

US-09-835-992A-19 (1-714) x Q17820 (1-343)

QY 646 CAGAAATAAATTAGAACAGTAA 623
|||||

Db 23 GlnAsnLysIleArgAsnSerLys 30

RESULT 9

Q8TTZ5 PRELIMINARY; PRT; 352 AA.

AC Q8TTZ5; 01-JUN-2002 (TREMblrel. 21, Created)

DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

DE Sulfate/molybdate ABC transporter, solute-binding protein.

GN MA0280.

OS Methanosarcina acetivorans.

OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;

OC Methanosarcinaceae; Methanosarcina.

OX NCBI_TaxID=2214;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C2A / ATCC 35395 / DSM 2834;

RX MEDLINE=21929760; PubMed=11932238;

RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,

RA Fitzhugh W., Calvo S., Engels S., Smirnov S., Atncor D., Brown A.,

RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,

RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,

RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,

RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,

RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,

RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,

RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,

RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,

RA Metcalf W.W., Birren B.;

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic

RT and physiological diversity.";

RL Genome Res. 12:532-542(2002).

DR EMBL; AE010687; AAM03733.1; -.

KW Complete proteome.

SQ SEQUENCE 352 AA; 38715 MW; 15652724987E8A9E CRC64;

Alignment Scores:

Pred. No.:	28.3	Length:	352
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	17	Gaps:	0

US-09-835-992A-19 (1-714) x Q8TTZ5 (1-352)

QY 114 TTTTCTAGTTTCTAATACTA 91
|||||

Db 10 PhePheLeuValPheLeuIleLeu 17

RESULT 10

Q8XN39 PRELIMINARY; PRT; 375 AA.

AC Q8XN39; 01-MAR-2002 (TREMblrel. 20, Created)

DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

DE Probable hexosyltransferase.

GN CPE0499.

OS Clostridium perfringens.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;

OC Clostridiales; Clostridiaceae; Clostridium.

OX NCBI_TaxID=1502;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=13 / TYPE A;

RX PubMed=11792842;

RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,

RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;

RT "Complete genome sequence of Clostridium perfringens, an anaerobic

RT flesh-eater.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).

DR EMBL; AP003187; BAB80205.1; -.

DR InterPro; IPR001296; Glycos_transf_1.

DR Pfam; PF00534; Glycos_transf_1; 1.

KW Transferase; Complete proteome.

SQ SEQUENCE 375 AA; 43028 MW; AB7626E3FF9E7B56 CRC64;

Alignment Scores:

Pred. No.:	27.9	Length:	375
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	16	Gaps:	0

US-09-835-992A-19 (1-714) x Q8XN39 (1-375)

QY 428 AGTGAGATTAAATCTAAATTT 405
|||||

Db 47 SerGlulIeLysAsnLeuLysIle 54

RESULT 11

Q92RV3 PRELIMINARY; PRT; 376 AA.

AC Q92RV3;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Putative permealase protein.

GN R00738 OR SMC00793.

OS Rhizobium melliloti (Sinorhizobium melliloti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RX MEDLINE=21396507; PubMed=11481430;

RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J., Boletard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S., Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D., Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U., Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;

RA "Analysis of the chromosome sequence of the legume symbiont Sinorhizobium melliloti strain 1021.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

DR EMBL; AL591784; CAC45310.1;

DR InterPro; IPR000560; HISAc-phsptase.

DR InterPro; IPR002549; UPF0118.

DR Pfam; PF01594; UPF0118; 1.

DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.

KW Complete proteome.

SO SEQUENCE 376 AA; 40673 MW; F119EE2DFB8CE87 CRC64;

Alignment Scores:

Pred. No.:	27.9	Length:	376
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	16	Gaps:	0

US-09-835-992A-19 (1-714) x Q92RV3 (1-376)

QY 493 CTGTTGGCATGTGTTCTTACTT 516
|||||

Db 97 LeuLeuAlaLeuCySpheLeuLeu 104

RESULT 12

Q972W9 PRELIMINARY; PRT; 413 AA.

AC Q972W9;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Putative D-hydantoinase.

GN ST1020.

OS Sulfolobus tokodaii.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI_TaxID=111955;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JCM 10545 / 7;

RX PubMed=11572479;

RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankal A., Kosugi H., Hosoyama A., Fukui S., Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y., Yoshizawa T., Tanaka T., Kuoh Y., Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Masuda S., Yanagi M., Nishimura M., Yamagishi A., Oshima T., Kikuchi H.;

RA "Complete genome sequence of an aerobic thermacidophilic

RT Crenarchaeon, Sulfolobus tokodaii strain7.";

RL DNA Res. 8:123-140(2001).

DR EMBL; AP000984; BAB66044.1;

DR InterPro; IPR002195; Dihydroorotase.

DR Pfam; PF00744; Dihydroorotase; 1.

KW Hypothetical protein; Complete proteome.

SO SEQUENCE 413 AA; 46760 MW; CE5B584E39710A9F CRC64;

Alignment Scores:

Pred. No.:	27.3	Length:	413
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	17	Gaps:	0

US-09-835-992A-19 (1-714) x Q972W9 (1-413)

QY 517 AGTTCCTCCCAAGGGAACCTCTTA 540
|||||

Db 405 SerSerProLysGlyLysLeuLeu 412

RESULT 13

Q9JM15 PRELIMINARY; PRT; 485 AA.

AC Q9JM15;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Neuronal glutamine transporter.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY;

RX MEDLINE=20127882; PubMed=10660562;

RA Varoqui H., Zhu H., Yao D., Ming H., Erickson J.D.;

RT "Cloning and Functional Identification of a Neuronal Glutamine Transporter.";

RL J. Biol. Chem. 275:4049-4054(2000).

DR EMBL; AF075704; AAF34240.1;

DR InterPro; IPR002422; AA/rel_primease2.

DR Pfam; PF01490; Aa-trans; 1.

SO SEQUENCE 485 AA; 53846 MW; 0CDAD864513FC186 CRC64;

Alignment Scores:

Pred. No.:	26.4	Length:	485
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	11	Gaps:	0

US-09-835-992A-19 (1-714) x Q9JM15 (1-485)

QY 618 GGGATTCTACTGTCTAATTGTA 641
|||||

Db 100 GlyTleuLeuPheLeuIleLeu 107

RESULT 14

Q99PR1 PRELIMINARY; PRT; 485 AA.

AC Q99PR1;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Amino acid transporter Nat-2.

GN AA408026 OR NAT.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21316561; PubMed=11325958;
RA Gu S., Roderick H.L., Camacho P., Jiang J.X.;
RT "Characterization of an N-system Amino Acid Transporter Expressed in
RT Retina and Its Involvement in Glutamine Transport.";
RL J. Biol. Chem. 276:24137-24144(2001).
DR EMBL; AF184240; AAG43433.2; -.
DR MGD; MGI:2145895; AA408026.
DR InterPro; IPR002422; AA/rel_primease2.
DR Pfam; PF01490; Aa_trans; 1.
SQ SEQUENCE 485 AA; 53739 MW; 11636A57030F976B CRC64;

Alignment Scores:

Pred. No.:	26.4	Length:	485
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	11	Gaps:	0

US-09-835-992A-19 (1-714) x Q99PR1 (1-485)

OY 618 GGGATTTTACTGTTCTAATTTTA 641
DB 100 GlyIleLeuPheLeuIleu 107

RESULT 15

O9HVA7 PRELIMINARY; PRT; 512 AA.

AC O9HVA7;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Iron (III)-transport system permease HltB.
GN HltB OR PA4688.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- FUNCTION: PROBABLY PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT
CC SYSTEM. PROBABLY RESPONSIBLE FOR THE TRANSLOCATION OF THE
CC SUBSTRATE ACROSS THE MEMBRANE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS.
DR EMBL; AE004882; AAG08075.1; -.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp; 2.
DR PROSITE; PS00402; BPD_TRANS_P_INN_MEMBER; 1.
KW Transmembrane; Transport; Complete proteome.
SQ SEQUENCE 512 AA; 56176 MW; 9A21A90B3CDFBF0A CRC64;

Alignment Scores:

Pred. No.:	26	Length:	512
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match: 3.67% Indels: 0
DB: 16 Gaps: 0

US-09-835-992A-19 (1-714) x Q9HVA7 (1-512)

OY 493 CTGTTGGCATGTGTTCTTACTT 516
DB 230 LeuLeuAlaLeuCysPheLeuLeu 237

Search completed: January 14, 2003, 17:36:49
Job time : 54.2755 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:24 ; Search time 33.8408 Seconds
(without alignments)
5410.226 Million cell updates/sec

Title: US-09-835-992A-20
Perfect score: 212
Sequence: 1 attttaatttctataaac.....cttggtatgcacagtgtgt 687

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09835992/runal_14012003_161525_3979/app_query.fasta_1.3932
-DB=A_Geneseq_101002 -OFMT=fastan -SUFFIX=olig.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdl
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992_@cgn_1_1_121_@runal_14012003_161525_3979 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YCAPOP=60 -YCAPEXT=60 -DELOP=6 -DELEXT=7

Database : A_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
C 1	8	3.8	29	19	AAW40006	Peptide effecting
2	8	3.8	104	22	AAO11927	Human polypeptide
3	8	3.8	159	22	ABB31093	Peptide #3744 enco
4	8	3.8	159	22	ABB36288	Peptide #3794 enco
5	8	3.8	159	22	ABB21653	Protein #3652 enco
6	8	3.8	159	22	AAW57060	Human brain expres
7	8	3.8	159	22	AAW69450	Human bone marrow
8	8	3.8	159	22	AAW17284	Peptide #3718 enco
9	8	3.8	159	22	AAW29784	Peptide #3821 enco
10	8	3.8	159	22	AAW04971	Peptide #3653 enco
11	8	3.8	159	23	ABG39070	Human peptide enco
12	8	3.8	275	22	ABB09061	Thermus caldophilu
13	8	3.8	342	13	AAW25835	Human platelet act
C 14	8	3.8	342	20	AAW49550	Human platelet act
C 15	8	3.8	342	22	ABB56381	Non-endogenous hum
C 16	8	3.8	342	23	ABB77904	Amino acid sequenc
C 17	8	3.8	342	23	ABB77918	Amino acid sequenc
18	8	3.8	542	22	AAW70385	Yeast host cell pr
C 19	8	3.8	631	21	AAW56847	Arabidopsis thailia
C 20	8	3.8	634	21	AAW56845	Arabidopsis thailia
C 21	8	3.8	743	21	AAW56845	Arabidopsis thailia
22	8	3.8	2870	21	AAW95559	Caenorhabditis ele
23	8	3.8	3178	21	AAW95556	Caenorhabditis ele
24	7	3.3	15	23	ABG62591	Eubacterial DNA po
25	7	3.3	18	10	AAW92063	Modified type of e
26	7	3.3	18	11	AAW05416	Improved hen egg w
27	7	3.3	19	19	AAW74938	Human secreted pro
28	7	3.3	21	9	AAW80973	Sequence of modifi
29	7	3.3	21	10	AAW92065	N-terminal region
30	7	3.3	21	11	AAW05417	Signal peptide.
31	7	3.3	22	20	AAW00284	Human secreted pro
32	7	3.3	43	21	AAW57883	Arabidopsis thailia
33	7	3.3	44	20	AAW14432	Human secreted pro
34	7	3.3	46	22	AAW22521	Novel human colon
35	7	3.3	46	22	AAW92476	Human digestive sy
36	7	3.3	47	22	AAW02178	Human polypeptide
C 37	7	3.3	48	22	ABB95983	Human testicular a
38	7	3.3	48	22	AAW95286	Human reproductive
C 39	7	3.3	52	21	AAW00502	Human secreted pro
40	7	3.3	55	20	AAW19543	Amino acid sequenc
41	7	3.3	57	22	ABBA1022	Peptide #8528 enco
42	7	3.3	57	22	ABB21506	Protein #7105 enco
43	7	3.3	57	22	AAW61879	Human brain expres
44	7	3.3	57	22	AAW74681	Human bone marrow
45	7	3.3	57	22	AAW20378	Peptide #6812 enco

ALIGNMENTS

RESULT 1
AAW40006
ID AAW40006 standard; peptide: 29 AA.
XX
AC AAW40006;
XX
DT 18-JUN-1998 (first entry)
XX
DE Peptide effecting G-protein-coupled receptor activity.
XX
KW G-protein-coupled receptor; GPCR; transmembrane domain; oligomerisation;
KW therapeutic composition; GPCR function; receptor monomeric form;
KW multimeric form; inhibition; GPCR-mediated process; GPCR binding;
KW treatment; disease; platelet activating factor.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9800538-A2.

XX 08-JAN-1998.
PD
XX
PF 01-JUL-1997; 97WO-IB00814.
XX
PR 01-JUL-1996; 96US-0021031.
XX
PA (BIOS-) BIOSIGNAL INC.
XX (UYMO-) UNIV MONTREAL.
PI Bouvier M, Dennis M, Hebert TE;
XX WPI; 1998-086964/08.
DR
XX
PT Peptide(s) or peptide leads affecting G protein-coupled receptor
PT activity - by altering receptor oligomerisation, useful in, e.g.
PT selectively modulating receptor function and treating neurological
PT or genetic diseases
XX
PS Claim 20; Page 62; 75pp; English.
XX
CC Peptides AAW40002-08 are modelled on transmembrane domains one to seven,
CC respectively, of human platelet activating factor, a G-protein-coupled
CC receptor (GPCR) (sic) whose activity is affected by the formation of
CC oligomers. GPCRs have a recurring pattern unique for the transmembrane
CC domains. The peptides are characterised by the ability to selectively
CC affect oligomerisation of the GPCR from which it was designed. The
CC peptides may be used with suitable carriers in the preparation of
CC therapeutic compositions. They may be administered to selectively
CC modulate GPCR function by affecting the ratio of receptor monomeric
CC to multimeric forms. The compositions can be administered to inhibit
CC GPCR-mediated processes by modulating GPCR binding to its agonist or
CC antagonist. The peptides may also be used to prevent or treat diseases
CC involving GPCRs. They are also useful for in vitro and in vivo studies
CC of GPCRs.
XX
SQ Sequence 29 AA;
XX
Alignment Scores:
Pred. No.: 32.4 Length: 29
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 19 Gaps: 0
US-09-835-992A-20 (1-687) x AAW40006 (1-29)
QY 103 TCTTTTCCCTAGTTTCTATA 80
Db 11 SerpHeLeuValPheLeuIle 18
RESULT 2
AAO11927
ID AAO11927 standard; Protein; 104 AA.
XX
AC AAO11927;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 25819.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX

PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI91858.
XX
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 20; SEQ ID NO 25819; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 104 AA;
XX
Alignment Scores:
Pred. No.: 28.8 Length: 104
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 22 Gaps: 0
US-09-835-992A-20 (1-687) x AAO11927 (1-104)
QY 351 TTGGGAAGAAATACTGTTCTG 374
Db 15 PheGLYLysIleLeuPheLeu 22
RESULT 3
ABB31093
ID ABB31093 standard; Peptide; 159 AA.
XX
AC ABB31093;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #3744 encoded by breast cell single exon nucleic acid probe.
XX
KW Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR


```

PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes
XX
XX Claim 27; SEQ ID NO 14061; 327pp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting
XX the probes with a collection of detectably labelled nucleic acids
XX derived from mRNA of human breast, and then measuring the label
XX bound to each probe of the microarray. The probes are useful for
XX verifying the expression of regions of genomic DNA predicted to
XX encode proteins. They are useful for gene discovery, and for
XX determining predisposition and/or prognosing breast disease. Gene
XX expression analysis is useful for assessing the toxicity of chemical
XX agents on cells. The microarray of this invention presents a far greater
XX diversity of probes for measuring gene expression, with far less bias
XX than expressed sequence tag microarrays. The method is suitable for
XX rapid production of functional information from genomic sequence. The
XX present sequence is a peptide encoded by a single exon nucleic acid
XX probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 159 AA;
XX
XX Alignment Scores:
XX Pred. No.: 27.7 Length: 159
XX Score: 8.00 Matches: 8
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 3.77% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-835-992A-20 (1-687) x ABB31093 (1-159)
XX
XX QY 479 CTGTTGGCATTTGTGTTCTTACTT 502
XX ||||||||||||||||||||||||
XX Db 120 LeuLeuAlaLeuCySpheLeuLeu 127
XX
XX RESULT 4
XX ID ABB36288 standard; Peptide; 159 AA.
XX
XX AC ABB36288;
XX
XX DT 04-FEB-2002 (first entry)
XX
XX DE Peptide #3794 encoded by human foetal liver single exon probe.
XX
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
XX OS Homo sapiens.
XX
XX PN WO200157277-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00669.
XX

```

PR	04-FEB-2000;	2000US-0180312.
PR	26-MAY-2000;	2000US-0207456.
PR	30-JUN-2000;	2000US-0608408.
PR	03-AUG-2000;	2000US-0632366.
PR	21-SEP-2000;	2000US-0234687.
PR	27-SEP-2000;	2000US-0236359.
PR	04-OCT-2000;	2000GB-0024263.
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX		
DR	WPI; 2001-483447/52.	
XX		
PT	Human genome-derived single exon nucleic acid probes useful for	
PT	analyzing gene expression in human fetal liver -	
XX		
PS	Claim 27; SEQ ID NO 28923; 639pp + sequence listing; English.	
XX		
CC	The invention relates to a single exon nucleic acid probe for	
CC	measuring human gene expression in a sample derived from human foetal	
CC	liver. The single exon nucleic acid probes may be used for predicting,	
CC	measuring and displaying gene expression in samples derived from human	
CC	fetal liver. The present sequence is a peptide encoded by a single exon	
CC	nucleic acid probe of the invention.	
CC	Note: The sequence data for this patent did not form part of the	
CC	printed specification, but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 159 AA;	
	Alignment Scores:	
	Pred. No.:	27.7
	Score:	8.00
	Percent Similarity:	100.00%
	Best Local Similarity:	100.00%
	Query Match:	3.77%
	DB:	22
		Gaps: 0
US-09-835-992A-20 (1-687) x ABB36288 (1-159)		
OY	479 CTGTTGGCATTTGTTTCCTTACTT	502
Db	120 LeuLeuAlaLeuCysPheLeuLeu	127
	RESULT 5	
	ABB21653	
ID	ABB21653 standard; protein; 159 AA.	
XX		
AC	ABB21653;	
XX		
DT	23-JAN-2002 (first entry)	
XX		
DE	Protein #3652 encoded by probe for measuring heart cell gene expression	
XX		
KW	Human; gene expression; heart; microarray; vascular system;	
KW	cardiovascular disease; hypertension; cardiac arrhythmia;	
KW	congenital heart disease.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157274-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	30-JAN-2001; 2001WO-US00666.	
XX		
PR	04-FEB-2000; 2000US-0180312.	
PR	26-MAY-2000; 2000US-0207456.	
PR	30-JUN-2000; 2000US-0608408.	
PR	03-AUG-2000; 2000US-0632366.	
PR	21-SEP-2000; 2000US-0234687.	
PR	27-SEP-2000; 2000US-0236359.	
PR	04-OCT-2000; 2000US-0234687.	
PR	27-SEP-2000; 2000US-0236359.	

PR 04-OCT-2000; 2000GB-0024263.
XX
PI (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
DR
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 15; SEQ ID NO 23423; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA1305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 159 AA;

Alignment Scores:
Pred. No.: 27.7 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-20 (1-687) x ABB21653 (1-159)

QY 479 CTGTTGCATGTGTTCTTACTT 502
XXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 120 LeuLeuAlaLeuCysPheLeuLeu 127

RESULT 6
AAM57060
ID AAM57060 standard; Protein; 159 AA.
XX
AC AAM57060;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29165.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
DR
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 29165; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 159 AA;

Alignment Scores:
Pred. No.: 27.7 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-20 (1-687) x AAM57060 (1-159)

QY 479 CTGTTGCATGTGTTCTTACTT 502
XXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 120 LeuLeuAlaLeuCysPheLeuLeu 127

RESULT 7
AAM69450
ID AAM69450 standard; Protein; 159 AA.
XX
AC AAM69450;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 29756.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
DR
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 29756; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.

SQ Sequence 159 AA;

Alignment Scores:

Pred. No.:	27.7	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	22	Gaps:	0

US-09-835-992a-20 (1-687) x AAM69450 (1-159)

OY 479 CTGTTGCATGTGTTCTTACTT 502
ID |||||
Db 120 LeuLeuAlaLeuCySpheLeu 127

RESULT 8
AAM17284
ID AAM17284 standard; Protein; 159 AA.
AC AAM17284;
XX 12-OCT-2001 (first entry)
DT
XX Peptide #3718 encoded by probe for measuring cervical gene expression.
DE
XX Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
XX Homo sapiens.
OS
XX WO200157278-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00670.
PE
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488901/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
PT
XX Claim 27; SEQ ID No 22110; 487bp; English.
PS
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 159 AA;

Alignment Scores:

Pred. No.:	27.7	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	22	Gaps:	0

US-09-835-992a-20 (1-687) x AAM17284 (1-159)

OY 479 CTGTTGCATGTGTTCTTACTT 502
ID |||||
Db 120 LeuLeuAlaLeuCySpheLeu 127

RESULT 9
AAM29784
ID AAM29784 standard; Protein; 159 AA.
XX
XX AAM29784;
AC
XX 17-OCT-2001 (first entry)
DT
XX Peptide #3821 encoded by probe for measuring placental gene expression.
DE
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
XX Homo sapiens.
OS
XX WO200157272-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00663.
PE
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-48897/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
PT
XX Claim 27; SEQ ID No 30053; 654bp; English.
PS
XX The present invention relates to single exon nucleic acid probes (SENPs:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
CC
SQ Sequence 159 AA;

Alignment Scores:

Pred. No.:	27.7	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match: 3.77% Indels: 0
DB: 22 Gaps: 0
US-09-835-992A-20 (1-687) x AAM04971 (1-159)
QY 479 CTGTTGGCATGTGTTCTTACTT 502
Db 120 LeuLeuAlaLeuCysPheLeuLeu 127
RESULT 10
AAM04971
ID AAM04971 standard; Protein; 159 AA.
XX
AC AAM04971;
XX
DT 09-OCT-2001 (first entry)
XX
DE Peptide #3653 encoded by probe for measuring breast gene expression.
XX
KW Probe; human; breast disease; breast cancer; development disorder;
KW Inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -
XX
PS Claim 27; SEQ ID NO 13711; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes
CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 159 AA;

Alignment Scores:
Pred. No.: 27.7 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-20 (1-687) x AAM04971 (1-159)
QY 479 CTGTTGGCATGTGTTCTTACTT 502
Db 120 LeuLeuAlaLeuCysPheLeuLeu 127
RESULT 11
ABG39070
ID ABG39070 standard; Peptide; 159 AA.
XX
AC ABG39070;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 28735.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
PS Claim 27; SEQ ID NO 28735; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 159 AA;

Alignment Scores:
Pred. No.: 27.7 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 23 Gaps: 0

US-09-835-992A-20 (1-687) x ABG39070 (1-159)

OY 479 CTGTTGGCATTGTGTTCTTACTT 502
Db 120 LeuLeuAlaLeuCysPheLeu 127

RESULT 12
ABB09061

ID ABB09061 standard; Protein; 275 AA.

XX AC ABB09061;

XX DT 26-JUN-2002 (first entry)

XX DE Thermus caldophilus GK24 lactose permease SEQ ID NO:13.

XX KW Thermus caldophilus GK24; heat resistant enzyme; cosmetic; food;

XX KM DNA polymerase; aqualysin I; beta-galactosidase; lactose permease.

XX OS Thermus caldophilus.

XX PN KR2001019888-A.

XX PD 15-MAR-2001.

XX PF 31-AUG-1999; 99KR-0036564.

XX PR 31-AUG-1999; 99KR-0036564.

XX PA (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.

XX PI Cho SJ, Jung BW, Kim HJ, Kim JS, Kim MG, Kim SY, Kim YS, Ko JH;

XX PI Ko SH, Lee DS, Lee JS, Park JA, Park JH, Shin HJ;

XX DR WPI; 2001-512441/56.

XX PT Base sequence and amino acid sequence of cosmid clone 3 isolated from
XX thermus caldophilus -

XX PS Example 2; Page 109-111; 130pp; Korean.

XX The present invention describes the nucleotide sequence of the cosmid

CC clone 3 isolated from Thermus caldophilus GK24, as given in the present
CC invention. Proteins encoded by the cosmid clone 3 are heat resistant
CC enzymes which can be used in the fields of cosmetic and food industries.
CC Thermus caldophilus GK24 has the optimal growth temperature of 72 plus
CC degrees Celsius and produces heat resistant enzymes including DNA
CC polymerase, aqualysin I, and beta-galactosidase. The DNA polymerase can
CC effectively be used in PCR which is performed at high temperature. The
CC present sequence represents lactose permease encoded by the Thermus
CC caldophilus GK24 cosmid clone 3.
XX
SQ Sequence 275 AA;

Alignment Scores:
Pred. No.: 26.3 Length: 275
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-20 (1-687) x ABB09061 (1-275)

OY 594 GTTCTTGGATTCTTACTGTTCTTA 617
Db 196 ValLeuTrpIleLeuLeuPheLeu 203

RESULT 13
AAR25835

ID AAR25835 standard; Protein; 342 AA.

XX AC AAR25835;

XX DT 21-JAN-1993 (first entry)

XX DE Human platelet activating factor receptor.

XX KW PAF; coagulation; blood clotting; clone phPAF.

XX OS Homo sapiens.

XX PN WO9212244-A.

XX PD 23-JUL-1992.

XX PF 27-DEC-1991; 91WO-JP01779.

XX PR 27-DEC-1990; 90JP-0407119.

XX PR 27-DEC-1990; 90JP-0407945.

XX PA (NISR) JAPAN TOBACCO INC.

XX PI Honda Z, Nakamura M, Shimizu T;

XX DR WPI; 1992-268667/32.

XX DR N-PSDB; AAQ26960.

XX PT DNA coding for human or guinea pig platelet activating factor
XX receptor - for prodn. of recombinant PAF receptor for screening
XX potential PAF agonists-antagonists

XX PS Claim 3; Fig 5; 50pp; Japanese.

XX The cDNA coding for human PAF receptor was isolated in a 1780 bp
XX clone designated phPAF. The clone consisted of a 112bp 5'-UTR (see
XX AAQ26961), a 1029bp ORF (AAQ26960) and a 639bp 3'-UTR (see AAQ26962).
XX The cDNA can be used to transform a suitable host cell line (e.g. COS
XX cells) for expression of the PAF receptor protein. The receptor
XX itself may be useful in screening of potential PAF agonists/
XX antagonists and in assays for PAF. See also AAQ26957-Q26959.

SQ Sequence 342 AA;

Alignment Scores:

Pred. No.: 25.8 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 13 Gaps: 0

US-09-835-992A-20 (1-687) x AAR25835 (1-342)

OY 103 TCTTTTTCCTAGTTTCTAATA 80
Db 194 SerPhePheValPheLeuIle 201

RESULT 14

AAV49550
ID AAV49550 standard; Protein; 342 AA.

AC AAV49550;

DT 13-JAN-2000 (first entry)

DE Human platelet activating factor receptor protein sequence.

KW Human; coding sequence polymorphism; vascular pathology gene;
KW polymorphic site; phenotype correlation; forensic; paternity testing;
KW medicine; genetic analysis; vascular disease.

OS Homo sapiens.

PN WO9950454-A2.

PD 07-OCT-1999.

PF 26-MAR-1999; 99WO-US06473.

PR 01-APR-1998; 98US-0054272.

PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

PI Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;

DR WPI; 1999-620066/53.

DR N-PSDB; AAZ32159.

PT Determination of polymorphisms in genes, especially those identifying
PT predisposition to vascular disease -

PS Disclosure; Fig 3; 134pp; English.

CC AAZ32159 to AAZ32194 represent reference alleles for specifically
CC claimed nucleic acid sequences from the present invention which comprise
CC polymorphic sites as given in a table in the specification, selected
CC from 92 single nucleotide polymorphisms in which the nucleotide at the
CC polymorphic site is different from a nucleotide at the same site in a
CC reference allele. The nucleic acids, and primers and probes, are used to
CC identify polymorphisms, which may predispose an individual to disease,
CC especially a vascular disease. They can also be used in phenotype
CC correlations, forensics, paternity testing, medicine or genetic
CC analysis. AAV49550 to AAV49573 represent the proteins which correspond
CC to some of the reference alleles.

SQ Sequence 342 AA;

Alignment Scores:

Pred. No.: 25.8 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 20 Gaps: 0

US-09-835-992A-20 (1-687) x AAV49550 (1-342)

OY 103 TCTTTTTCCTAGTTTCTAATA 80
Db 194 SerPhePheValPheLeuIle 201

RESULT 15

ABB56381
ID ABB56381 standard; Protein; 342 AA.

AC ABB56381;

DT 18-FEB-2002 (first entry)

DE Non-endogenous human GPCR protein, SEQ ID NO: 555.

KW Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
KW constitutively activated GPCR; agonist; disease.

OS Homo sapiens.
OS Synthetic.

PN WO200177172-A2.

PD 18-OCT-2001.

PF 05-APR-2001; 2001WO-US11098.

PR 07-APR-2000; 2000US-195747P.

PA (AREN-) ARENA PHARM INC.

PI Lehmann-Brulnsma K, Liaw CW, Lin I;

DR WPI; 2001-648759/74.

DR N-PSDB; ABI98017.

PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in
PT disease treatment, comprises contacting candidate compounds with
PT versions of GPCRs -

PS Claim 1; Page 356-357; 394pp; English.

CC The invention relates to G protein-coupled receptors (GPCRs) for which
CC the endogenous ligand has been identified. Non-endogenous
CC constitutively activated versions of known GPCRs are used in the
CC invention for the direct identification of candidate compounds as
CC receptor agonists, inverse agonists or partial agonists. Such
CC agonists are useful as therapeutic agents for diseases or disorders
CC associated with GPCRs. The present sequence is a non-endogenous
CC version of a known human GPCR.

SQ Sequence 342 AA;

Alignment Scores:

Pred. No.: 25.8 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-20 (1-687) x ABB56381 (1-342)

OY 103 TCTTTTTCCTAGTTTCTAATA 80

Db 194 SerPhePheValPheLeuIle 201

Search completed: January 14, 2003, 17:27:19
Job time : 35.8408 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:54 ; Search time 9.44148 Seconds
(without alignments)
4281.862 Million cell updates/sec

Title: US-09-835-992A-20

Perfect score: 212

Sequence: 1 attttaattctctataaac.....cttggtatgcacagtgtgt 687

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+n2p.model -DEV-xlp
-O=/cgn2_1/USPTO_spool/US09835992/runat_14012003_161527_4050/app_query.fasta_1.3932
-DB-Issued_Patents_AA -QFMT-fastan -SUFFIX-olig.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=-1 -MATRIX-oligo -TRANS-human40.cdl
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992@cgn2_1_1_49@runat_14012003_161527_4050 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	542	1	US-08-246-583-3
2	7	3.3	18	6	5182195-2
3	7	3.3	19	4	US-09-149-476-534
4	7	3.3	21	6	5182195-7
5	7	3.3	44	4	US-09-288-143-88
6	7	3.3	62	1	US-08-488-961-6
7	7	3.3	62	4	US-08-973-297-6
8	7	3.3	62	5	PCT-US96-06511-6
9	7	3.3	71	4	US-09-134-001C-3466
10	7	3.3	187	4	US-09-149-476-417
11	7	3.3	222	1	US-07-869-933-12
12	7	3.3	222	1	US-07-869-933-28

13	7	3.3	222	4	US-09-103-663-12	Sequence 12, Appl
14	7	3.3	222	4	US-09-103-663-28	Sequence 28, Appl
15	7	3.3	267	2	US-08-557-128-4	Sequence 4, Appl1
16	7	3.3	267	4	US-09-242-690A-36	Sequence 36, Appl
17	7	3.3	326	1	US-07-603-133B-25	Sequence 25, Appl
18	7	3.3	380	4	US-08-857-076-110	Sequence 110, App
19	7	3.3	427	4	US-09-134-001C-5143	Sequence 5143, Ap
20	7	3.3	527	1	US-08-246-583-2	Sequence 2, Appl1
21	7	3.3	538	2	US-08-933-227-5	Sequence 5, Appl1
22	7	3.3	554	4	US-09-321-276-4	Sequence 4, Appl1
23	7	3.3	554	4	US-08-916-481-2	Sequence 2, Appl1
24	7	3.3	563	4	US-08-916-481-3	Sequence 3, Appl1
25	7	3.3	663	4	US-09-196-293-5	Sequence 5, Appl1
26	7	3.3	663	4	US-08-209-603E-5	Sequence 5, Appl1
27	7	3.3	663	4	US-08-235-836C-70	Sequence 70, Appl
28	7	3.3	663	4	US-08-235-836C-78	Sequence 78, Appl
29	7	3.3	681	4	US-08-760-615-4	Sequence 4, Appl1
30	7	3.3	681	4	US-08-760-615-6	Sequence 6, Appl1
31	7	3.3	700	4	US-08-235-836C-74	Sequence 74, Appl
32	7	3.3	1698	4	US-09-315-793-12	Sequence 12, Appl
33	7	3.3	1724	4	US-08-857-076-12	Sequence 12, Appl
34	7	3.3	1956	4	US-08-843-417-2	Sequence 2, Appl1
35	7	3.3	1957	4	US-08-669-656A-2	Sequence 2, Appl1
36	7	3.3	1957	4	US-08-669-656A-8	Sequence 8, Appl1
37	7	3.3	2132	4	US-08-669-656A-6	Sequence 6, Appl1
38	6	2.8	10	1	US-08-197-793-3	Sequence 3, Appl1
39	6	2.8	10	2	US-08-636-176-3	Sequence 3, Appl1
40	6	2.8	10	5	PCT-US95-01618-3	Sequence 3, Appl1
41	6	2.8	11	1	US-08-338-634-4	Sequence 4, Appl1
42	6	2.8	11	1	US-08-787-547-67	Sequence 67, Appl
43	6	2.8	11	5	PCT-US95-16415-1	Sequence 1, Appl1
44	6	2.8	14	1	US-08-232-453A-40	Sequence 40, Appl
45	6	2.8	15	4	US-09-027-900-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-246-583-3
Sequence 3, Application US/08246583
Patent No. 5750394
GENERAL INFORMATION:
APPLICANT: Palese, Peter
TITLE OF INVENTION: IDENTIFICATION AND USE OF ANTIVIRAL
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION OF HOST CELL PROTEINS
TITLE OF INVENTION: AND VIRAL PROTEINS REQUIRED FOR VIRAL REPLICATION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/246,583
FILING DATE: 20-MAY-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 6923-040
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-246-583-3

Alignment Scores:
Pred. No.: 4.97 Length: 542
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-20 (1-687) x US-08-246-583-3 (1-542)

OY 497 TTACTTAGTCTCTCCCAAGAAAC 520
    ||| ||| ||| ||| ||| ||| |||
Db 349 LeuLeuSerProlySGluasn 356

RESULT 2
5182195-2
; Patent No. 5182195
; APPLICANT: NAKAHAMA, KAZUO;KAISHO, YOSHIHIKO;YOSHIMURA, KOJI
; TITLE OF INVENTION: METHOD FOR INCREASING USING PROTEASE
; DEFICIENT YEASTS
; NUMBER OF SEQUENCES: 71
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/269,140
; FILING DATE: 09-NOV-1988
; SEQ ID NO:2:
; LENGTH: 18
5182195-2

Alignment Scores:
Pred. No.: 80.7 Length: 18
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 6 Gaps: 0

US-09-835-992A-20 (1-687) x 5182195-2 (1-18)

OY 479 CTGTTGGCATGTGTTCTTA 499
    ||| ||| ||| ||| ||| ||| |||
Db 6 LeuLeuAlaLeuCysPheLeu 12

RESULT 3
US-09-149-476-534
; Sequence 534, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; EARLIER FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
```

```

; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
```

EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501

EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Alignment Scores:

Pred. No.: 80.1 Length: 19
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 4 Gaps: 0

US-09-835-992a-20 (1-687) x US-09-149-476-534 (1-19)

QY 606 TTACTGTCCTAATTTATTC 626
Db 4 LeuLeuPheLeuIleLeuPhe 10

RESULT 4

5182195-7

Patent No. 5182195
APPLICANT: NAKAHAMA, KAZUO;KAISHO, YOSHIHIKO;YOSHIMURA, KOJI
TITLE OF INVENTION: METHOD FOR INCREASING USING PROTEASE
DEFICIENT YEASTS

NUMBER OF SEQUENCES: 71

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/269,140

FILING DATE: 09-NOV-1988

SEQ ID NO:7:

LENGTH: 21

5182195-7

Alignment Scores:

Pred. No.: 79 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 6 Gaps: 0

US-09-835-992a-20 (1-687) x 5182195-7 (1-21)

QY 479 CTGTTGGCATTGTCTTCTTA 499

|||||
Db 6 leuLeuAlaLeuCysPheLeu 12

RESULT 5
US-09-288-143-88
; Sequence 88, Application US/09288143
; Patent No. 6433139
; GENERAL INFORMATION:
; APPLICANT: Brewer et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P1
; CURRENT APPLICATION NUMBER: US/09/288,143
; EARLIER FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/US98/21142
; EARLIER FILING DATE: 1998-10-08
; EARLIER APPLICATION NUMBER: 60/061,463
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,529
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/071,498
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,527
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,536
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,532
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals stop translation
US-09-288-143-88

Alignment Scores:
Pred. No.: 71.1 Length: 44
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-20 (1-687) x US-09-288-143-88 (1-44)

QY 603 ATTTACTGTTCCTAATTTTA 623
Db 5 lleLeuLeuPheLeuIleLeu 11

RESULT 6
US-08-488-961-6
; Sequence 6, Application US/08488961
; Patent No. 5606042
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel S.
; APPLICANT: Walker, John C.
; TITLE OF INVENTION: Glycine and Phaseolus
; TITLE OF INVENTION: alpha-D-Galactosidases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P. O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099-4390
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,961
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: P-320 (Umo)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 689-3500
TELEFAX: (810) 689-4071
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 62 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-488-961-6

Alignment Scores:
Pred. No.: 67.7 Length: 62
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-20 (1-687) x US-08-488-961-6 (1-62)

QY 482 TTGCATGTGTTCTTACTT 502
Db 20 leuAlaLeuCysPheLeu 26

RESULT 7
US-08-973-297-6
; Sequence 6, Application US/08973297
; Patent No. 6184017
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel S.
; APPLICANT: Walker, John C.
; TITLE OF INVENTION: Glycine and Phaseolus
; TITLE OF INVENTION: alpha-D-Galactosidases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 NO. 6184017thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,297
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 0994.00050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 539-5050
; TELEFAX: (810) 539-5055
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 62 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
US-08-973-297-6

Alignment Scores:

Pred. No.:	67.7	Length:	62
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	4	Gaps:	0

US-09-835-992A-20 (1-687) x US-08-973-297-6 (1-62)

OY 482 TTGGCATGTGTTCTTACTT 502
|||||

Db 20 LeuAlaLeuCySpheLeu 26

RESULT 8
PCT-US96-06511-6

Sequence 6, Application PC/TUS9606511

GENERAL INFORMATION:

APPLICANT: Smith, Daniel S.

APPLICANT: Walker, John C.

TITLE OF INVENTION: Glycine and Phaseolus

TITLE OF INVENTION: alpha-D-galactosidases

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kohn & Associates

STREET: 30500 Northwestern Hwy., Suite 410

CITY: Farmington Hills

STATE: Michigan

COUNTRY: US

ZIP: 48334

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/06511

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Kohn, Kenneth I.

REGISTRATION NUMBER: 30,955

REFERENCE/DOCKET NUMBER: 0994.00050

TELECOMMUNICATION INFORMATION:

TELEPHONE: (810) 539-5050

TELEFAX: (810) 539-5055

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 62 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

PCT-US96-06511-6

Alignment Scores:

Pred. No.:	67.7	Length:	62
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	5	Gaps:	0

US-09-835-992A-20 (1-687) x PCT-US96-06511-6 (1-62)

OY 482 TTGGCATGTGTTCTTACTT 502
|||||

Db 20 LeuAlaLeuCySpheLeu 26

RESULT 9
US-09-134-001C-3466

Sequence 3466, Application US/09134001C

Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 3466

LENGTH: 71

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3466

Alignment Scores:

Pred. No.:	66.4	Length:	71
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	4	Gaps:	0

US-09-835-992A-20 (1-687) x US-09-134-001C-3466 (1-71)

OY 240 CCATCATTAAGATAATTGAA 220
|||||

Db 13 ProSerLeuArgIleIleGlu 19

RESULT 10
US-09-149-476-417

Sequence 417, Application US/09149476

Patent No. 6420526

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002P1

CURRENT APPLICATION NUMBER: US/09/149,476

CURRENT FILING DATE: 1998-09-08

EARLIER APPLICATION NUMBER: PCT/US98/04493

EARLIER FILING DATE: 1998-03-06

EARLIER APPLICATION NUMBER: 60/040,162

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,333

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/038,621

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,626

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,334

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,336

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,163

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/047,600

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,615

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,597

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,502

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,633

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,583

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,617

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Alignment Scores:

Pred. No.:	57.9	Length:	187
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	4	Gaps:	0

US-09-835-992a-20 (1-687) x US-09-149-476-417 (1-187)

QY 150 TTGTATTAATTTCTTTT 130

Db 6 PheValIleIlePheLeuphe 12

RESULT 11

US-07-869-933-12

Sequence 12, Application US/07869933

Patent No. 5770396

GENERAL INFORMATION:

APPLICANT: KINET, Jean-Pierre

TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF

TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR

TITLE OF INVENTION: IMMUNOGLOBULIN

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/869,933

FILING DATE: 19920416

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 40399/154 NIHD

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 222 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Rat
STRAIN: ECRI alpha subunit
US-07-869-933-12

Alignment Scores:

Pred. No.:	56.5	Length:	222
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	1	Gaps:	0

US-09-835-992a-20 (1-687) x US-07-869-933-12 (1-222)

QY 69 AATAACATTAAGTATGAAA 89

Db 134 AsnAsnIleSerIleArgLys 140

RESULT 12

US-07-869-933-28

Sequence 28, Application US/07869933

Patent No. 5770396

GENERAL INFORMATION:

APPLICANT: KINET, Jean-Pierre

TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF

TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR

TITLE OF INVENTION: IMMUNOGLOBULIN

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/869,933

FILING DATE: 19920416

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 40399/154 NIHD

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 222 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

STRAIN: alpha subunit

US-07-869-933-28

Alignment Scores:

Pred. No.:	56.5	Length:	222
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	1	Gaps:	0

US-09-835-992a-20 (1-687) x US-07-869-933-28 (1-222)

QY 69 AATAACATTAAGTATGAAA 89

Db 134 AsnAsnIleSerIleArgLys 140

RESULT 13

US-09-103-663-12
; Sequence 12, Application US/09103663D
; Patent No. 6171803
; GENERAL INFORMATION:
; APPLICANT: Kinet et al.
; TITLE OF INVENTION: Isolation, characterization, and use of the human beta
; TITLE OF INVENTION: subunit of the high affinity receptor for
; TITLE OF INVENTION: immunoglobulin E.
; FILE REFERENCE: 50490
; CURRENT APPLICATION NUMBER: US/09/103,663D
; CURRENT FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 07/869,933
; EARLIER FILING DATE: 1992-04-16
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-103-663-12

Alignment Scores:

Pred. No.:	56.5	Length:	222
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	4	Gaps:	0

US-09-835-992A-20 (1-687) x US-09-103-663-12 (1-222)

OY 69 AATAACATTAGTATTAGAAAA 89

Db 134 AsnAsnIleSerIleArgLys 140

RESULT 14

US-09-103-663-28
; Sequence 28, Application US/09103663D
; Patent No. 6171803
; GENERAL INFORMATION:
; APPLICANT: Kinet et al.
; TITLE OF INVENTION: Isolation, characterization, and use of the human beta
; TITLE OF INVENTION: subunit of the high affinity receptor for
; TITLE OF INVENTION: immunoglobulin E.
; FILE REFERENCE: 50490
; CURRENT APPLICATION NUMBER: US/09/103,663D
; CURRENT FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 07/869,933
; EARLIER FILING DATE: 1992-04-16
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-103-663-28

Alignment Scores:

Pred. No.:	56.5	Length:	222
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	4	Gaps:	0

US-09-835-992A-20 (1-687) x US-09-103-663-28 (1-222)

OY 69 AATAACATTAGTATTAGAAAA 89

|||||

Db 134 AsnAsnIleSerIleArgLys 140

RESULT 15

US-08-557-128-4
; Sequence 4, Application US/08557128
; Patent No. 5849524
; GENERAL INFORMATION:
; APPLICANT: KONDO, Koji
; APPLICANT: KAJIWARA, Susumu
; APPLICANT: MISAWA, NO. 5849524hiko
; TITLE OF INVENTION: TRANSFORMATION SYSTEMS FOR THE YEAST
; TITLE OF INVENTION: CANDIDA UTILIS AND THE EXPRESSION OF HETEROLOGOUS GENES
; TITLE OF INVENTION: THEREWITH
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,128
; FILING DATE: 25-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP95/01005
; FILING DATE: 25-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-129287
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-285823
; FILING DATE: 26-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-135015
; FILING DATE: 25-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-557-128-4

Alignment Scores:

Pred. No.:	55	Length:	267
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-20 (1-687) x US-08-557-128-4 (1-267)

OY 507 GAACCTAGTAGAACAACAAT 487

Db 80 GluLeuSerLysLysHisAsn 86

|||||

Wed Jan 15 08:41:06 2003

us-09-835-992a-20.olig.ra

Search completed: January 14, 2003, 17:42:09
Job time : 10.6915 secs



THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:29:25 ; Search time 6.57721 Seconds
(without alignments)
4052.943 Million cell updates/sec

Title: US-09-835-992A-20
Perfect score: 212
Sequence: 1 attttaatttctataaac.....cttggtatgcacagttgtt 687

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 118974 seqs, 19401057 residues

Word size: 1

Total number of hits satisfying chosen parameters: 214280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+np.model -DEV=xlp
-O=/cgn2_1/USPRO_pool/US09835992/runat_14012003_161528_4118/app_query.fasta_1.3932
-DB=Published_Applications_AA -QFMT=fastlan -SUFFIX=olig.rapb -MINMATCH=0.1
-LOOPL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09835992.ecgn_1_1_28@runat_14012003_161528_4118
-NCPU=6 -ICPU=3 -NO_XLPRY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	159	10	US-09-864-761-36951
2	8	3.8	327	12	US-10-084-206-3
3	7	3.3	57	10	US-09-864-761-40404
4	7	3.3	81	9	US-10-002-344A-243

5	7	3.3	127	10	US-09-764-877-1466	Sequence 1466, Ap
6	7	3.3	197	10	US-09-780-717-47	Sequence 47, Appl
7	7	3.3	241	10	US-09-740-668A-26	Sequence 26, Appl
8	7	3.3	241	12	US-10-059-964-62	Sequence 62, Appl
9	7	3.3	267	10	US-09-908-855-36	Sequence 36, Appl
10	7	3.3	334	10	US-09-815-242-13249	Sequence 13249, A
11	7	3.3	380	10	US-09-205-658-110	Sequence 110, App
12	7	3.3	380	10	US-09-844-353A-110	Sequence 110, App
13	7	3.3	481	10	US-09-815-242-10888	Sequence 10888, A
14	7	3.3	554	10	US-09-800-396-4	Sequence 4, Appl1
15	7	3.3	758	10	US-09-925-299-859	Sequence 859, App
16	7	3.3	1724	10	US-09-205-658-12	Sequence 12, Appl
17	7	3.3	1724	10	US-09-844-353A-12	Sequence 12, Appl
18	6	2.8	9	10	US-09-834-765-444	Sequence 444, App
19	6	2.8	10	8	US-08-452-843A-13	Sequence 13, Appl
20	6	2.8	10	10	US-09-834-765-82	Sequence 82, Appl
21	6	2.8	10	10	US-09-834-765-611	Sequence 611, App
22	6	2.8	11	9	US-09-909-460-67	Sequence 67, Appl
23	6	2.8	17	10	US-09-949-196-4	Sequence 4, Appl1
24	6	2.8	20	10	US-09-742-732-5	Sequence 5, Appl1
25	6	2.8	22	9	US-10-001-876-130	Sequence 130, App
26	6	2.8	22	10	US-09-949-196-32	Sequence 32, Appl
27	6	2.8	22	10	US-09-949-196-36	Sequence 36, Appl
28	6	2.8	22	10	US-09-949-196-40	Sequence 40, Appl
29	6	2.8	25	12	US-10-001-879-113	Sequence 113, App
30	6	2.8	27	10	US-09-864-761-39910	Sequence 39910, A
31	6	2.8	28	10	US-09-929-818-104	Sequence 104, App
32	6	2.8	30	10	US-09-071-838-227	Sequence 227, App
33	6	2.8	31	10	US-09-864-761-38920	Sequence 38920, A
34	6	2.8	32	10	US-09-864-761-47089	Sequence 47089, A
35	6	2.8	34	10	US-09-864-761-34176	Sequence 34176, A
36	6	2.8	36	10	US-09-864-761-47949	Sequence 47949, A
37	6	2.8	37	10	US-09-864-761-44137	Sequence 44137, A
38	6	2.8	39	10	US-09-764-847-937	Sequence 937, App
39	6	2.8	40	10	US-09-864-761-42251	Sequence 42251, A
40	6	2.8	41	10	US-09-864-761-46929	Sequence 46929, A
41	6	2.8	43	10	US-09-864-761-44190	Sequence 44190, A
42	6	2.8	43	10	US-09-764-847-866	Sequence 866, App
43	6	2.8	48	12	US-10-001-879-199	Sequence 199, App
44	6	2.8	49	9	US-09-843-676-147	Sequence 147, App
45	6	2.8	49	9	US-09-766-253-147	Sequence 147, App

ALIGNMENTS

RESULT 1
US-09-864-761-36951
Sequence 36951, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 36951
;; LENGTH: 159
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC007374.3
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
;; OTHER INFORMATION: EST_HUMAN HIT: AA485663.1, EVALUATE 1.00e-67
US-09-864-761-36951

Alignment Scores:
Pred. No.: 5.64 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-20 (1-687) x US-09-864-761-36951 (1-159)
QY 479 CTGTTGCATGTGTTCTTACTT 502
|||||
Db 120 LeuLeuAlaLeuCysPheLeuLeu 127

RESULT 2
US-10-084-206-3
;; Sequence 3, Application US/10084206
;; Patent No. US20020106741A1
;; GENERAL INFORMATION:
;; APPLICANT: Li et al.
;; TITLE OF INVENTION: G-Protein Receptor HTNAD29
;; FILE REFERENCE: PF191D1C1
;; CURRENT APPLICATION NUMBER: US/10/084,206
;; PRIOR FILING DATE: 2002-02-28
;; PRIOR APPLICATION NUMBER: US95/07288
;; PRIOR FILING DATE: 1995-06-06
;; PRIOR APPLICATION NUMBER: 08/468,534
;; PRIOR FILING DATE: 1995-06-06
;; PRIOR APPLICATION NUMBER: 09/399,095
;; PRIOR FILING DATE: 1999-09-20

;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: PatentIn Version 3.1
;; SEQ ID NO 3
;; LENGTH: 327
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-084-206-3

Alignment Scores:
Pred. No.: 4.97 Length: 327
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 12 Gaps: 0

US-09-835-992A-20 (1-687) x US-10-084-206-3 (1-327)
QY 103 TCTTTTCTAGTTTCTATA 80
|||||
Db 190 SerPhePheLeuAlaPheLeuLeu 197

RESULT 3
US-09-864-761-40404
;; Sequence 40404, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;; FILE REFERENCE: Aeonica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; PRIOR FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40404
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL078591.18
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
; OTHER INFORMATION: SWISSPROT HIT: Q26563, EVALUATE 2.70e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE169919.1, EVALUATE 6.20e+00
US-09-864-761-40404

Alignment Scores:
Pred. No.: 62.9 Length: 57
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-20 (1-687) x US-09-864-761-40404 (1-57)

QY 626 CTGAACCTCCATTTCCTCCCA 646

Db 15 LeuLySLeuHISphSeSerPro 21

RESULT 4
US-10-002-344A-243
; Sequence 243, Application US/10002344A
; Patent No. US20020172959A1
; GENERAL INFORMATION:
; APPLICANT: Recipon, Herve
; APPLICANT: Sun, Yongming
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0241
; CURRENT APPLICATION NUMBER: US/10/002,344A
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/242,998
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 243
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo saplen
US-10-002-344A-243

Alignment Scores:
Pred. No.: 59.1 Length: 81
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 9 Gaps: 0

US-09-835-992A-20 (1-687) x US-10-002-344A-243 (1-81)

QY 603 ATTTTACTGTCTCTAATTTA 623

Db 73 IleLeuLeuPheLeuIleLeu 79

RESULT 5
US-09-764-877-1466

; Sequence 1466, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; PRIOR APPLICATION data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1466
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1466

Alignment Scores:
Pred. No.: 54.6 Length: 127
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-20 (1-687) x US-09-764-877-1466 (1-127)

QY 312 GCCTGTGTCAAGCTCTGATC 332

Db 80 AlacysvalGlnAlaLeuIle 86

RESULT 6
US-09-780-717-47
; Sequence 47, Application US/09780717
; Patent No. US20010044941A1
; GENERAL INFORMATION:
; APPLICANT: Helentjaris, Tim
; APPLICANT: Bates, Nic
; APPLICANT: Allen, Stephen M.
; TITLE OF INVENTION: of use
; TITLE OF INVENTION: of use
; FILE REFERENCE: 035718/208677
; CURRENT APPLICATION NUMBER: US/09/780,717
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,509
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Glycine max
US-09-780-717-47

Alignment Scores:
Pred. No.: 50.5 Length: 197
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-20 (1-687) x US-09-780-717-47 (1-197)

QY 29 CAAAGCATTTATTTATCTAT 49

Db 45 GlnSerIleIleLeuSerTyr 51

RESULT 7

US-09-740-668A-26
; Sequence 26, Application US/09740668A
; Patent No. US20020076700A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; TITLE OF INVENTION: No. US20020076700A1 polypeptides and nucleic acids encoding sam
; FILE REFERENCE: 15966-537 CIP
; CURRENT APPLICATION NUMBER: US/09/740,668A
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: PCT/US99/29584
; PRIOR FILING DATE: 1999-12-17
; PRIOR APPLICATION NUMBER: 09/465,512
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113,485
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/112,837
; PRIOR FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)..(241)
; OTHER INFORMATION: wherein xaa is any amino acid as defined in the specification.
US-09-740-668A-26

Alignment Scores:

Pred. No.:	48.8	Length:	241
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	10	Gaps:	0

US-09-835-992A-20 (1-687) x US-09-740-668A-26 (1-241)

OY 324 GCTCTGATCATATTCTTTTA 344
|||||
Db 126 Alaleullellepheleuleu 132

RESULT 8

US-10-059-964-62
; Sequence 62, Application US/10059964
; Patent No. US20020120115A1
; GENERAL INFORMATION:
; APPLICANT: Rikihisa, Yasuko
; APPLICANT: Ohasi, No. US20020120115A110
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; TITLE OF INVENTION: Chafeensis
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/10/059,964
; CURRENT FILING DATE: 2002-01-28
; EARLIER APPLICATION NUMBER: 09/314,701
; EARLIER FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 241
; TYPE: PRT
; ORGANISM: P30-12
US-10-059-964-62

Alignment Scores:

Pred. No.:	48.8	Length:	241
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	12	Gaps:	0

US-09-835-992A-20 (1-687) x US-10-059-964-62 (1-241)
OY 500 CTTAGTCTCCAGAAAC 520
|||||
Db 110 LeuSerSerProlySGluAsn 116

RESULT 9

US-09-908-855-36
; Sequence 36, Application US/09908855
; Patent No. US20020115220A1
; GENERAL INFORMATION:
; APPLICANT: KONDO, KEIJI
; APPLICANT: MIURA, YUTAKA
; TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 049441/0118
; CURRENT APPLICATION NUMBER: US/09/908,855
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 09/242,690
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: JP 8/241062
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Candida utilis
US-09-908-855-36

Alignment Scores:

Pred. No.:	47.9	Length:	267
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	10	Gaps:	0

US-09-835-992A-20 (1-687) x US-09-908-855-36 (1-267)

OY 507 GAAGTAAGTAAACACAAAT 487
|||||
Db 80 GluLeuSerLysLysHisAsn 86

RESULT 10

US-09-815-242-13249
; Sequence 13249, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931


```
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13249
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13249

Alignment Scores:
Pred. No.: 46.1          Length: 334
Score: 7.00             Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32%      Indels: 0
DB: 10                  Gaps: 0

US-09-835-992a-20 (1-687) x US-09-815-242-13249 (1-334)

OY 388 GCATTTCATGCTATCAGAAC 368
Db 121 AlaphelisaIaIeArgAn 127

RESULT 11
US-09-205-658-110
; Sequence 110, Application US/09205658
; Patent NO. US20010029617A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/205,658
; CURRENT FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: 08/857,076
; EARLIER FILING DATE: 1997-05-15
; EARLIER APPLICATION NUMBER: 08/888,534
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: US98/10080
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-205-658-110

Alignment Scores:
Pred. No.: 45          Length: 380
Score: 7.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32%      Indels: 0
DB: 10                Gaps: 0

US-09-835-992a-20 (1-687) x US-09-205-658-110 (1-380)

OY 94 CTAGTTTCTAATACTAATG 74
Db 4 LeuValPheLeuIleLeuMet 10

RESULT 12
US-09-844-353a-110
; Sequence 110, Application US/09844353A
; Patent NO. US20020037585A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutarou
; APPLICANT: Patterson, Garth

APPLICANT: Ogg, Scott
APPLICANT: Paradis, Suzanne
APPLICANT: Tissenbaum, Heidi
APPLICANT: Morris, Jason
APPLICANT: Kowek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: 00786/351005
; CURRENT APPLICATION NUMBER: US/09/844,353A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 08/857,076
; PRIOR FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-844-353a-110

Alignment Scores:
Pred. No.: 45          Length: 380
Score: 7.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32%      Indels: 0
DB: 10                Gaps: 0

US-09-835-992a-20 (1-687) x US-09-844-353a-110 (1-380)

OY 94 CTAGTTTCTAATACTAATG 74
Db 4 LeuValPheLeuIleLeuMet 10

RESULT 13
US-09-815-242-10888
; Sequence 10888, Application US/09815242
; Patent NO. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10888
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10888
```

Alignment Scores:

Pred. No.:	43.2	Length:	481
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	10	Gaps:	0

US-09-835-992A-20 (1-687) x US-09-815-242-10888 (1-481)

OY 191 AAACGTCATCTGAAGCAA 211

Db 399 LysThrAlaSerGlnSerLys 405

RESULT 14

US-09-800-396-4

; Sequence 4, Application US/09800396
; Patent No. US20020065395A1

; GENERAL INFORMATION:

; APPLICANT: Wallis, Nicola

; TITLE OF INVENTION: NOVEL RESPONSE REGULATOR

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESS: Dechert Price & Rhoads

; STREET: 997 Lenox Drive, Building 3, Suite 210

; CITY: Lawrenceville

; STATE: NJ

; COUNTRY: USA

; ZIP: 08543

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/800,396

; FILING DATE: 06-Mar-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/879,531

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Bloom, Allen

; REGISTRATION NUMBER: 29,135

; REFERENCE/DOCKET NUMBER: GM10018

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 609-520-3214

; TELEFAX: 609-520-3259

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 554 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-800-396-4

US-09-800-396-4

Alignment Scores:

Pred. No.:	42.1	Length:	554
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	10	Gaps:	0

US-09-835-992A-20 (1-687) x US-09-800-396-4 (1-554)

OY 97 TTCTAGTTTCTAATACTA 77

Db 6 PheLeuValPheLeuIleLeu 12

RESULT 15

US-09-925-299-859

; Sequence 859, Application US/09925299

; Patent No. US20020055627A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA102

; CURRENT APPLICATION NUMBER: US/09/925,299

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05883

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; NUMBER OF SEQ ID NOS: 1556

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 859

; LENGTH: 758

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (590)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-299-859

Alignment Scores:

Pred. No.:	39.9	Length:	758
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	10	Gaps:	0

US-09-835-992A-20 (1-687) x US-09-925-299-859 (1-758)

OY 372 GAACAGTATTCTTCCCA 352

Db 151 GluThrValPheSerSerGln 157

Search completed: January 14, 2003, 18:08:18
Job time : 9.57721 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:45 ; Search time 23.0202 Seconds
(without alignments)
5737.940 Million cell updates/sec

Title: US-09-835-992A-20
Perfect score: 212
Sequence: 1 attttaatttctataaac.....cttggtatgcacagtgtt 687

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565536

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:
-MODEL-frame+_n2p.model -DEV-xlp
-O-/cgn2_1/USPRO_spool/US09835992/runat_14012003_161526_4038/app_query.fasta_1.3932
-DB-PIR_73 -OPMT-fastan -SUFFIX-oligo.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER-US09835992@cgn_1_1_192_@runat_14012003_161526_4038 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YCAPOP=60 -YCAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	268	2 F64024	hypothetical prote
2	8	3.8	342	2 A40191	platelet-activatin
3	8	3.8	343	2 T19082	hypothetical prote
4	8	3.8	393	2 H64233	hypothetical prote
5	8	3.8	512	2 E83060	iron (III)-transpo
6	8	3.8	542	2 S30884	SRP1 protein - Yea
7	8	3.8	791	2 H72552	hypothetical prote
8	8	3.8	1183	2 F90559	conserved hypothet
9	8	3.8	2329	2 T28125	hypothetical prote
10	7	3.3	55	2 G95972	hypothetical prote
11	7	3.3	65	2 S34538	hypothetical prote
12	7	3.3	91	2 T11269	NADH2 dehydrogenas
13	7	3.3	105	2 E84075	hypothetical prote
14	7	3.3	120	2 T30769	hypothetical prote

15	7	3.3	126	2 F72150	B6L protein - vari
16	7	3.3	164	1 NWMU1	2S albumin 1 precu
17	7	3.3	164	1 NFMU3	2S albumin 3 precu
18	7	3.3	194	2 C69933	negative regulatio
19	7	3.3	195	2 E71000	hypothetical prote
20	7	3.3	201	2 G90674	probable transcrip
21	7	3.3	201	2 B85525	probable transcrip
22	7	3.3	201	2 F64758	yahd protein - Esc
23	7	3.3	205	2 T80398	calmodulin homolog
24	7	3.3	207	2 T24267	hypothetical prote
25	7	3.3	236	1 B64212	hypothetical prote
26	7	3.3	245	2 A30154	hypothetical prote
27	7	3.3	254	2 D69140	hypothetical prote
28	7	3.3	269	2 AF1532	hypothetical prote
29	7	3.3	269	2 AH1175	spermidine/putresc
30	7	3.3	271	2 F72379	sugar ABC transpor
31	7	3.3	276	2 AC1586	gamma-glutamyl kin
32	7	3.3	276	2 AD1232	gamma-glutamyl kin
33	7	3.3	278	2 C75616	oxidoreductase, sh
34	7	3.3	279	2 B75442	3-hydroxybutyryl-C
35	7	3.3	279	2 T20443	hypothetical prote
36	7	3.3	290	2 G72203	sugar ABC transpor
37	7	3.3	307	2 G64350	quinolinate synthe
38	7	3.3	313	2 D86829	ribose ABC transpo
39	7	3.3	314	2 T45010	probable site-spec
40	7	3.3	314	2 S73717	hypothetical prote
41	7	3.3	319	2 D90589	hypothetical prote
42	7	3.3	320	2 T30894	lipase homolog T0
43	7	3.3	326	1 VGXRHH	glycoprotein VP7 p
44	7	3.3	326	1 VGXRHR	glycoprotein VP7 p
45	7	3.3	326	1 A44891	glycoprotein VP7 p

ALIGNMENTS

RESULT 1
F64024
hypothetical protein H11273 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
C:Accession: F64024
R.Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:9550630; PMID:7542800
A:Accession: F64024
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-268 <TIGR>
A:Cross-references: GB:U32807; GB:L42023; NID:g1574725; PIDN:ACC22921.1; PID:g1574728

Alignment Scores:

Pred. No.: 10
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.79%
DB: 2
Gaps: 0

US-09-835-992A-20 (1-687) x F64024 (1-268)

OY 522 GAGTTTCCTTGAGAGACTAAGT 499
Db 223 Glupheserleuglycylulenser 230

RESULT 2
A40191
platelet-activating factor receptor - human
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 20-Jun-2000

C;Accession: A40191; JH0479; A41079; JC1359; A42831; I51923
R;Kunz, D.; Gerard, N.P.; Gerard, C.
J. Biol. Chem. 267, 9101-9106, 1992
A;Title: The human leukocyte platelet-activating factor receptor. cDNA cloning, cell su
A;Reference number: A40191; MUID:92250505; PMID:1374385
A;Accession: A40191
A;Molecule type: mRNA
A;Residues: 1-342 <KUN>
A;Cross-references: GB:M76674; NID:9456293; PIDN:AAA60002.1; PID:9456294
R;Ye, R.D.; Prossnitz, E.R.; Zou, A.; Cochran, C.G.
Biochem. Biophys. Res. Commun. 180, 105-111, 1991
A;Title: Characterization of a human cDNA that encodes a functional receptor for platele
A;Reference number: JH0479; MUID:92028922; PMID:1656963
A;Accession: JH0479
A;Molecule type: mRNA
A;Residues: 1-342 <YER>
A;Cross-references: GB:M80436; NID:9189537; PIDN:AAA60001.1; PID:9189538
A;Experimental source: granulocyte, cell line HL-60 all
R;Nakamura, M.; Honda, Z.; Izumi, T.; Sakanaka, C.; Mutoh, H.; Minami, M.; Bito, H.; Sey
J. Biol. Chem. 266, 20400-20405, 1991
A;Title: Molecular cloning and expression of platelet-activating factor receptor from hu
A;Reference number: A41079; MUID:92041873; PMID:1657923
A;Accession: A41079
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-342 <NAK>
A;Cross-references: GB:D10202; GB:D90433; NID:9219975; PIDN:BAA01050.1; PID:9219976
R;Sugimoto, T.; Tsuchimoto, H.; McGregor, C.G.A.; Mutoh, H.; Shimizu, T.; Kurachi, Y.
Biochem. Biophys. Res. Commun. 189, 617-624, 1992
A;Title: Molecular cloning and characterization of the platelet-activating factor recept
A;Reference number: JC1359; MUID:93112021; PMID:1281995
A;Accession: JC1359
A;Molecule type: mRNA
A;Residues: 1-315, 'N', 317-342 <SUG>
A;Experimental source: heart
A;Note: the authors translated the codon AAT for residue 316 as Lys
R;Seyfried, C.E.; Schweickart, V.L.; Godiska, R.; Gray, P.W.
Genomics 13, 832-834, 1992
A;Title: The human platelet-activating factor receptor gene (PTAFR) contains no introns
A;Reference number: A42831; MUID:92347886; PMID:1322356
A;Accession: A42831
A;Molecule type: DNA
A;Residues: 1-226, 'TG', 229-342 <SEY>
A;Cross-references: GB:M88177; NID:9190697; PIDN:AAA60214.1; PID:9190698
A;Note: sequence extracted from NCBI backbone (NCBIN:109813, NCBIP:109814)
R;Chase, P.B.; Halonen, M.; Regan, J.W.
Am. J. Respir. Cell Mol. Biol. 8, 240-244, 1993
A;Title: Cloning of a human platelet-activating factor receptor gene: evidence for an in
A;Reference number: I51923; MUID:93192035; PMID:8383507
A;Accession: I51923
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-342 <RES>
A;Cross-references: GB:S56396; NID:9298580; PIDN:AAB25755.1; PID:9298581
C;Genetics:
A;Gene: GDB:PTAFR
A;Cross-references: GDB:128806; OMIM:173393
A;Map position: 1p35-1p34.3
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein
F;17-38/Domain: transmembrane #status predicted <TRI>
F;54-75/Domain: transmembrane #status predicted <TII>
F;92-113/Domain: transmembrane #status predicted <III>
F;134-155/Domain: transmembrane #status predicted <TIV>
F;184-205/Domain: transmembrane #status predicted <TRV>
F;233-253/Domain: transmembrane #status predicted <TVI>
F;277-297/Domain: transmembrane #status predicted <VII>

Alignment Scores:
Pred. No.: 9.81 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.79% Indels: 0
DB: 2 Gaps: 0
US-09-835-992A-20 (1-687) x A40191 (1-342)
QY 103 TCTTTTTCCTAGTTTTCATA 80
Db 194 SerpHeLeuValPheLeuIle 201
RESULT 3
hypotheical protein C08B6.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T19082
R;Wilkinson, J.
submitted to the EMBL data library, May 1996
A;Reference number: Z19070
A;Accession: T19082
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-343 <WIL>
A;Cross-references: EMBL:Z72502; PIDN:CAA96591.1; GSPDB:GN00023; CESP:C08B6.9
A;Experimental source: clone C08B6
C;Genetics:
A;Gene: CESP:C08B6.9
A;Map position: 5
A;Introns: 25/2; 84/3; 174/2; 290/2

Alignment Scores:
Pred. No.: 9.8 Length: 343
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x T19082 (1-343)

QY 628 CAGATAAAATAGACAGTAA 605
Db 23 GlnAsnLysIleArgAsnSerLys 30

RESULT 4
H64233
hypotheical protein MG306 - Mycoplasma genitalium
C;Species: Mycoplasma genitalium
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 07-Dec-1999
C;Accession: H64233
R;Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,
, C.A.; Venter, J.C.
Science 270, 397-403, 1995
A;Title: The minimal gene complement of Mycoplasma genitalium.
A;Reference number: A64200; MUID:96026346; PMID:7569993
A;Accession: H64233
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-393 <TIGR>
A;Cross-references: GB:U39712; GB:L43967; NID:91046005; PID:91046006; TIGR:MG306
A;Experimental source: strain G-37
C;Genetics:
A;Genetic code: SGC3

Alignment Scores:
Pred. No.: 9.68 Length: 393
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x H64233 (1-393)

Oy 103 TCCTTTTCCTAGTTTCTAATA 80
Db 19 SerPhePheLeuValPheLeuIle 26

RESULT 5
E83060

Iron (III)-transport system permease hltB PA4688 [imported] - Pseudomonas aeruginosa (st
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83060
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Llm,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: E83060
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <STO>
A:Cross-references: GB:AE004882; GB:AE004091; NID:g9950939; PIDN:AAG08075.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: hltB; PA4688
C:Superfamily: sfub protein

Alignment Scores:

Pred. No.:	9.45	Length:	512
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-20 (1-687) x E83060 (1-512)

Oy 479 CTGTTGGCATTGTGTTCTTACTT 502
Db 230 LeuLeuAlaLeuCysPheLeuLeu 237

RESULT 6
S30884

SRP1 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein N1606; protein YNL189w
C:Species: Saccharomyces cerevisiae
C:Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 20-Jun-2000
C:Accession: S30884; S63144
R:Yano, R.; Oakes, M.; Yamagishi, M.; Dodd, J.A.; Nomura, M.
Mol. Cell. Biol. 12, 5640-5651, 1992
A:Title: Cloning and characterization of SRP1, a suppressor of temperature-sensitive RNA
A:Reference number: S30884; MUID:93078765; PMID:1448093
A:Accession: S30884
A:Molecule type: DNA
A:Residues: 1-542 <YAN>
A:Cross-references: EMBL:M75849; NID:g172702; PIDN:AAA35090.1; PID:g172703
R:Obermaier, B.; Piravandi, E.; Rinke, M.; Domdey, H.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63122
A:Accession: S63144
A:Molecule type: DNA
A:Residues: 1-542 <OHE>
A:Cross-references: EMBL:Z71465; NID:g1302178; PIDN:CAA96083.1; PID:g1302179; MIPS:YNL18
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:SRP1; SCM1; KAP1
A:Cross-references: SGD:S0005133; MIPS:YNL189w
A:Map position: 14L
C:Superfamily: pendulin
C:Keywords: nucleus; transmembrane protein
F:336-352/Domain: transmembrane #status predicted <TMM>

Alignment Scores:

Pred. No.:	9.4	Length:	542
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-20 (1-687) x S30884 (1-542)

Oy 497 TTACTTACTTCTCCCAAGAAAC 520
Db 349 LeuLeuSerSerProlySGLuAsn 356

RESULT 7
H72552

hypothetical protein APE1708 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: H72552
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: H72552
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-791 <KAW>
A:Cross-references: DDBJ:AP000062; NID:g5105244; PIDN:BAA80709.1; PID:d1044495; PID:g
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1708

Alignment Scores:

Pred. No.:	9.07	Length:	791
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-20 (1-687) x H72552 (1-791)

Oy 76 TTAGTATTAGAAAAGTAGGAAA 99
Db 134 LeuValLeuGluLysLeuGlyLys 141

RESULT 8
F90559

conserved hypothetical protein MYPU_3820 [imported] - Mycoplasma pulmonis (strain UAB
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: F90559
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer,
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: F90559
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1183 <KUR>
A:Cross-references: GB:AL445566; PID:g14089796; PIDN:CAC13555.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPU_3820
A:Genetic code: SGC3

Alignment Scores:

Pred. No.:	8.74	Length:	1183
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.79%	Indels:	0

DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x F90559 (1-1183)

OY 194 TTTCAATATATATCATTTCA 171

Db 577 PheserAsnTyraSnleuileSer 584

RESULT 9

T28125

hypothetical protein ZK945.9 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C/Accession: T28125

R/Wilkinson, J.

submitted to the EMBL Data Library, March 1995

A/Reference number: Z20472

A/Accession: T28125

A/Status: preliminary; translated from GB/EMBL/DDbJ

A/Molecule type: DNA

A/Residues: 1-2329 <WIL>

A/Cross-references: EMBL:Z48544; NID:g695493; PIDN:CAA8844.2.1; GSPDB:GN00020; CESP:ZK945

A/Experimental source: clone ZK945

C/Genetics:

A/Gene: CESP:ZK945.9

A/Map position: 2

A/Introns: 160/3; 187/1; 243/1; 580/3; 607/3; 705/3; 789/1; 864/3; 940/3; 974/3; 1064/2;

Alignment Scores:

Pred. No.: 8 21 Length: 2329

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.77% Indels: 0

DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x T28125 (1-2329)

OY 398 AGATTTTATCTCNCCTAATTGA 421

Db 1632 ArgPheLeuIleSerLeuIleLeu 1639

RESULT 10

G95972

hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid pSymH

C/Species: Sinorhizobium meliloti

C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C/Accession: G95972

R/Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A/Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A/Reference number: A95842; MUID:21396508; PMID:11481431

A/Accession: G95972

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-55 <KUR>

A/Cross-references: GB:AL591985; PIDN:CAC49447.1; PID:g15140933; GSPDB:GN00167

A/Experimental source: strain 1021, megaplasmid pSymB

R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A/Reference number: A96039; MUID:21368234; PMID:11474104

A/Accession: T11269

C/Contents: annotation

C/Genetics:

A/Gene: Smb21672

A/Genome: plasmid

Alignment Scores:

Pred. No.: 136 Length: 55

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.32% Indels: 0

DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x G95972 (1-55)

OY 529 AGTTAAGAGTTTCTTGGG 509

Db 25 SerLeuArgValPheLeuGly 31

RESULT 11

S34538

hypothetical protein 65 (rps18 3' region) - Euglena gracilis chloroplast

C/Species: chloroplast Euglena gracilis

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 29-Sep-1999

C/Accession: S34538; S34905

R/Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.; Monfort, A.; Orsat, B.; Spielma

submitted to the EMBL Data Library, January 1993

A/Description: The complete sequence of the Euglena gracilis chloroplast genome (tent

A/Reference number: S34494

A/Accession: S34538

A/Molecule type: DNA

A/Residues: 1-65 <HAL1>

A/Cross-references: EMBL:X70810; NID:g415327; PIDN:CAA50118.1; PID:g415774

R/Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.R.; Monfort, A.; Orsat, B.; Spiel

Nucleic Acids Res. 21, 3537-3544, 1993

A/Title: Complete sequence of Euglena gracilis chloroplast DNA.

A/Reference number: S34862; MUID:93347989; PMID:8346031

A/Accession: S34905

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-65 <HAL2>

A/Cross-references: EMBL:X70810; NID:g415327; PIDN:CAA50118.1; PID:g415774

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

C/Genetics:

A/Genome: chloroplast

C/Superfamily: conserved hypothetical protein ycf9

C/Keywords: chloroplast

Alignment Scores:

Pred. No.: 134 Length: 65

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.32% Indels: 0

DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x S34538 (1-65)

OY 100 TTTTCTAGTTTCTAATA 80

Db 51 PhePheLeuValPheLeuIle 57

RESULT 12

T11269

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - hardbacked tick (Ixodes hexa

C/Species: mitochondrion Ixodes hexagonus

C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002

C/Accession: T11269

R/Black IV, W.C.; Roehrdanz, R.L.

Mol. Biol. Evol. 15, 1772-1785, 1998

A/Title: Mitochondrial gene order is not conserved in arthropods: prostriate and meta

A/Reference number: Z17252; MUID:99083443; PMID:9866211

A/Accession: T11269

A/Status: preliminary; translated from GB/EMBL/DDbJ

A/Molecule type: DNA

A/Residues: 1-91 <BLA>

A/Cross-references: EMBL:AF081828; NID:g4164542; PID:g4164552; PIDN:AAD05513.1

C/Genetics:

A/Genome: mitochondrion

C/Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Alignment Scores:

Pred. No.:	130	Length:	91
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-20 (1-687) x T11269 (1-91)

OY 225 TTATTCCTTAATGATGCTGTTT 245

Db 37 LeuPheLeuMetMetValPhe 43

RESULT 13

E84075 hypothetical protein BH3405 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C;Accession: E84075

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: E84075

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-105 <STO>

A;Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BAB07124.1; GSPDB:GNOC

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH3405

Alignment Scores:

Pred. No.:	128	Length:	105
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-20 (1-687) x E84075 (1-105)

OY 606 TTAAGTCTTCTAATTTTATTC 626

Db 11 LeuLeuPheLeuIleLeuPhe 17

RESULT 14

T30769

hypothetical protein 8L - vaccinia virus (strain Ankara)

N;Alternate names: putative 13.7k protein

C;Species: vaccinia virus

A;Variety: strain Ankara

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T30769

R;Antoine, G.; Schelflinger, F.; Falkner, F.G.; Dorner, F.

submitted to the EMBL Data Library, March 1997

A;Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain

A;Reference number: Z20877

A;Accession: T30769

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-120 <ANT>

A;Cross-references: EMBL:U94848; PIDN:AAB96403.1

A;Experimental source: strain Ankara

C;Genetics:

A;Note: MVA008L

Alignment Scores:

Pred. No.:	126	Length:	120
Score:	7.00	Matches:	7

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 3.30%

DB: 2

US-09-835-992A-20 (1-687) x T30769 (1-120)

OY 360 AAAATACCTGTTCTGATAGCA 380

Db 2 LysIleLeuPheLeuIleAla 8

RESULT 15

F72150

B6L protein - variola minor virus (strain Garcia-1966)

C;Species: variola minor virus

C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999

C;Accession: F72150

R;Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Saifonov, P.F.; Massung, R.F.; Lo

submitted to GenBank, March 1998

A;Description: Analysis of the complete coding sequence of DNA of alastrim variola m

A;Reference number: A72150

A;Accession: F72150

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-126 <SHC>

A;Cross-references: GB:Y16780; NID:g5830555; PIDN:CAB54600.1; PID:g1542556; PID:g5830

A;Experimental source: strain Garcia-1966

C;Genetics:

A;Gene: B6L

Alignment Scores:

Pred. No.:	126	Length:	126
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-20 (1-687) x F72150 (1-126)

OY 360 AAAATACCTGTTCTGATAGCA 380

Db 2 LysIleLeuPheLeuIleAla 8

Search completed: January 14, 2003, 17:40:39
Job time : 27.0202 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:54 ; Search time 10.6084 Seconds
(without alignments)
5372.014 Million cell updates/sec

Title: US-09-835-992A-20
Perfect score: 212
Sequence: 1 attttaattttatttaaac.....cttgttatgcacagtgtt 687

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+_n2p.model -DEV-olp
-O=/cgn2_1/USPTO_spool/US09835992/runat_14012003_161525_4014/app_query.fasta_1.3932
-DB=SwissProt_40 -QFMT=fastan -SUFFIX=olig.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdl
-LIST=45 -DOCALLIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992_@CGN_1_1_77_@runat_14012003_161525_4014 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	8	3.8	208	1	PAFR_MACMU
C 2	8	3.8	268	1	YC73_HAEIN
C 3	8	3.8	342	1	PAFR_HUMAN
C 4	8	3.8	393	1	Y306_MYCGE
C 5	8	3.8	542	1	IMAI_YEAST
C 6	8	3.8	3178	1	YCF9_EUGER
C 7	7	3.3	65	1	YCF9_EUGER
C 8	7	3.3	164	1	2SS1_ARATH
C 9	7	3.3	164	1	2SS3_ARATH
C 10	7	3.3	180	1	YPKW_THECU
C 11	7	3.3	194	1	YPBH_BACSU
C 12	7	3.3	201	1	YAHB_ECOLI
C 13	7	3.3	245	1	FCE1_RAT
C 14	7	3.3	267	1	PYRF_PICST
C 15	7	3.3	268	1	PYRF_CANTR
C 16	7	3.3	276	1	PROB_LISIN
C 17	7	3.3	276	1	PROB_LISMO
C 18	7	3.3	278	1	HBD_DEIRA

C 19	7	3.3	278	1	Y110_MYCGE	P47356	mycoplasma
C 20	7	3.3	307	1	NADA_METJA	O57850	methanococc
C 21	7	3.3	314	1	Y315_MYCPN	O50362	mycoplasma
C 22	7	3.3	326	1	VS09_ROTET	O03874	equine rota
C 23	7	3.3	326	1	VS09_ROTTH	P11855	human rotav
C 24	7	3.3	326	1	VS09_ROTTH	P12476	rhesus rota
C 25	7	3.3	345	1	CLT2_PIG	O95n03	sus scrofa
C 26	7	3.3	372	1	FLGI_BUCAI	P57426	buchnera ap
C 27	7	3.3	379	1	PANE_YEAST	P38787	saccharomyc
C 28	7	3.3	380	1	CAPM_STAUA	P39862	staphylococ
C 29	7	3.3	405	1	DXR_PASMU	P57985	pasteurella
C 30	7	3.3	423	1	MTB5_NEIGO	O59605	neisseria g
C 31	7	3.3	423	1	MTN4_NEILA	P50182	neisseria l
C 32	7	3.3	473	1	IAIC_MALDO	P37821	malus domes
C 33	7	3.3	527	1	YO99_CAEEL	P41849	caenorhabdi
C 34	7	3.3	536	1	IMAS_HUMAN	O15131	homo sapien
C 35	7	3.3	538	1	IMAI_HUMAN	P52294	homo sapien
C 36	7	3.3	538	1	IMAI_MOUSE	O60960	mus musculu
C 37	7	3.3	551	1	VP17_YEAST	P32913	saccharomyc
C 38	7	3.3	560	1	YJK5_YEAST	P42948	saccharomyc
C 39	7	3.3	584	1	LEU1_SCHPO	O59736	schizosacch
C 40	7	3.3	613	1	UL06_EBV	P03213	epstein-bar
C 41	7	3.3	681	1	VGP_MABVM	P35253	marburg vir
C 42	7	3.3	688	1	YAOA_SCHPO	Q10109	schizosacch
C 43	7	3.3	690	1	SYGB_BUCAI	P57235	buchnera ap
C 44	7	3.3	778	1	RHP9_SCHPO	P87074	schizosacch
C 45	7	3.3	1068	1	YCF0_MARPO	P12221	marchantia

ALIGNMENTS

RESULT 1
ID PAFR_MACMU STANDARD; PRT; 208 AA.
AC P35366;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Platelet activating factor receptor (PAF-R) (Fragment).
GN PTAFR.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Behal R.H., Debussere M.S., Olson M.S.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR PLATELET ACTIVATING FACTOR, A CHEMOTACTIC
CC PHOSPHOLIPID MEDIATOR THAT POSSESSES POTENT INFLAMMATORY, SMOOTH-
CC MUSCLE CONTRACTILE AND HYPOTENSIVE ACTIVITY. SEEM TO MEDIATE ITS
CC ACTION VIA A G PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCINUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L07333; AAA68893.1;
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm_1; 1.
CC PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Chemotaxis.
FT DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 17 38 1 (POTENTIAL).
FT DOMAIN 39 54 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 55 74 2 (POTENTIAL).
FT DOMAIN 75 91 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 92 113 3 (POTENTIAL).
FT DOMAIN 114 133 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 134 155 4 (POTENTIAL).
FT DOMAIN 156 184 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 185 205 5 (POTENTIAL).
FT DOMAIN 206 >208 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 90 173 BY SIMILARITY.
FT NON_TER 208 208
SQ SEQUENCE 208 AA; 23736 MW; 7985928F70B3C6A1 CRC64;

Alignment Scores:

Pred. No.: 9.48 Length: 208
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-20 (1-687) x PAFR_MACMU (1-208)

OY 103 TCTTTTCTAGTTTCTAATA 80
DB 194 SerpPheLeuValPheLeuLe 201
|||||

RESULT 2

YC73_HAEIN STANDARD; PRT; 268 AA.
ID YC73_HAEIN
AC P44150;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein HI1273.

GN HI1273.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI_TaxID=727;

RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-Rd / KW20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

RA Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,

RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,

RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

RA Venter J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus influenzae

RT Rd.";

RL Science 269:496-512(1995).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: U32807; AAC22921.1; -
CC TIGR: HI1273; -

DR InterPro; IPR000051; SAM_bind.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 268 AA; 30510 MW; E5B28DA7AADC4D0B CRC64;

Alignment Scores:

Pred. No.: 9.03 Length: 268
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-20 (1-687) x YC73_HAEIN (1-268)

OY 522 GAGTTTCTTGAGAGACTAAGT 499
DB 223 GluPheSerLeuGlyGluLeuSer 230
|||||

RESULT 3

PAFR_HUMAN STANDARD; PRT; 342 AA.
ID PAFR_HUMAN
AC P25105;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Platelet activating factor receptor (PAF-R).

GN PAFR OR PAFR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=92028922; PubMed=1656963;

RA Ye R.D., Prossnitz E.R., Zou A., Cochran C.G.;

RT "Characterization of a human cDNA that encodes a functional receptor

RT for platelet activating factor.";

RL Biochem. Biophys. Res. Commun. 180:105-111(1991).

RN [2]
RP SEQUENCE FROM N.A.

RX MEDLINE=92041873; PubMed=1657923;

RA Nakamura M., Honda Z., Izumi T., Sakanaka C., Mutoh H., Minami M.,

RA Bito H., Seyama Y., Matsumoto T., Noma M., Shimizu T.;

RT "Molecular cloning and expression of platelet-activating factor

RT receptor from human leukocytes.";

RL J. Biol. Chem. 266:20400-20405(1991).

RN [3]
RP SEQUENCE FROM N.A.

RX MEDLINE=92250505; PubMed=1374385;

RA Kunz D., Gerard N.P., Gerard C.;

RT "The human leukocyte platelet-activating factor receptor. cDNA

RT cloning, cell surface expression, and construction of a novel

RT epitope-bearing analog.";

RL J. Biol. Chem. 267:9101-9106(1992).

RN [4]
RP SEQUENCE FROM N.A.

RX MEDLINE=92347886; PubMed=1322356;

RA Seyfried C.E., Schweickart V.L., Godiska R., Gray P.W.;

RT "The human platelet-activating factor receptor gene (PTAFR) contains

RT no introns and maps to chromosome 1.";

RL Genomics 13:832-834(1992).

RN [5]
RP SEQUENCE FROM N.A.

RC TISSUE=Heart;

RX MEDLINE=93112021; PubMed=1281995;

RA Sugimoto T., Tsuchimochi H., McGregor C.G., Mutoh H., Shimizu T.,

RA Kurachi Y.;

RT "Molecular cloning and characterization of the platelet-activating

RT factor receptor gene expressed in the human heart.";

RL Biochem. Biophys. Res. Commun. 189:617-624(1992).

RN [6]
RP SEQUENCE FROM N.A.

RA Behal R.H., Debussere M.S., Olson M.S.;

RL Submitted (xxx-1992) to the EMBL/GenBank/DBJ databases.

RN [7]
RP SEQUENCE FROM N.A.

RX MEDLINE-93192035; PubMed-8383507;
RA Chase P.B., Halonen M., Regan J.W.;
RT "Cloning of a human platelet-activating factor receptor gene:
evidence for an intron in the 5'-untranslated region.";
RL Am. J. Respir. Cell Mol. Biol. 8:240-244(1993).
CC -1- FUNCTION: RECEPTOR FOR PLATELET ACTIVATING FACTOR, A CHEMOTACTIC
CC PHOSPHOLIPID MEDIATOR THAT POSSESSES POTENT INFLAMMATORY, SMOOTH-
CC MUSCLE CONTRACTILE AND HYPOTENSIVE ACTIVITY. SEEM TO MEDIATE ITS
CC ACTION VIA A G PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M80436; AAA60001.1; -
DR EMBL; M76674; AAA60002.1; -
DR EMBL; D10202; BAA01050.1; -
DR EMBL; M88177; AAA60214.1; -
DR EMBL; S52624; AAB24695.2; -
DR EMBL; L07334; AAA60108.1; -
DR EMBL; S56396; AAB25755.1; -
DR PIR; JH0479; JH0479.
DR PIR; A40191; A40191.
DR PIR; A41079; A41079.
DR Genew; HGNC:9582; PTAFR.
DR MIM; 173393; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Chemotaxis;
KW Polymorphism.
FT DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 17 38 1 (POTENTIAL).
FT DOMAIN 39 54 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 55 74 2 (POTENTIAL).
FT DOMAIN 75 91 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 92 113 3 (POTENTIAL).
FT DOMAIN 114 133 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 134 155 4 (POTENTIAL).
FT DOMAIN 156 184 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 185 205 5 (POTENTIAL).
FT DOMAIN 206 233 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 234 254 6 (POTENTIAL).
FT DOMAIN 255 276 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 277 296 7 (POTENTIAL).
FT DOMAIN 297 342 CYTOPLASMIC (POTENTIAL).
FT DISULFID 90 173 BY SIMILARITY.
FT CARBOHYD 169 169 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 224 224 A -> D (IN DBSNP:5938).
FT VARIANT 224 224 /FTID-VAR_011851.
FT VARIANT 338 338 N -> S (IN DBSNP:5939).
FT VARIANT 338 338 /FTID-VAR_011852.
FT CONFLICT 28 28 L -> P (IN REF. 6).
FT CONFLICT 66 66 F -> L (IN REF. 6).
FT CONFLICT 95 95 C -> R (IN REF. 6).
FT CONFLICT 227 228 KR -> TG (IN REF. 4).
FT CONFLICT 227 228 KR -> TT (IN REF. 6).
FT CONFLICT 247 247 P -> A (IN REF. 6).
FT CONFLICT 316 316 K -> N (IN REF. 5).
SQ SEQUENCE 342 AA; 39203 MW; 890073C9EBA79228 CRC64;

Alignment Scores:
Pred. No.: 8.61 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 1 Gaps: 0
US-09-835-992a-20 (1-687) x PAFR_HUMAN (1-342)
OY 103 TCCTTTTCTAGTTTCTAATA 80
Db 194 SerpHeLeuValPheLeuIle 201
RESULT 4
Y306_MYCGE STANDARD; PRT; 393 AA.
ID Y306_MYCGE
AC P47548;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG306.
GN MG306.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE-96026346; PubMed-7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39711; AAC71528.1; -
DR TIGR; MG306; -
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 131 151 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 253 273 POTENTIAL.
FT TRANSMEM 289 309 POTENTIAL.
FT TRANSMEM 349 369 POTENTIAL.
SQ SEQUENCE 393 AA; 45750 MW; D740FDA979EC364A CRC64;

Alignment Scores:
Pred. No.: 8.39 Length: 393
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 1 Gaps: 0
US-09-835-992a-20 (1-687) x Y306_MYCGE (1-393)
OY 103 TCCTTTTCTAGTTTCTAATA 80
Db 19 SerpHeLeuValPheLeuIle 26
RESULT 5
IMAL_YEAST

ID IMA1_YEAST STANDARD; PRT; 542 AA.
AC Q02821;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Importin alpha subunit (Karyopherin alpha subunit) (Serine-rich RNA
polymerase I suppressor protein).
GN SRP1 OR KAP60 OR YNL189W OR N1606.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93078765; PubMed=1448093;
RA Yano R., Oakes M., Yamagishi M., Dodd J.A., Nomura M.;
RT "Cloning and characterization of SRP1, a suppressor of temperature-
sensitive RNA polymerase I mutations, in Saccharomyces cerevisiae.";
RL Mol. Cell. Biol. 12:5640-5651(1992).
RN [2]
RP SEQUENCE FROM N.A. AND FUNCTION.
RC STRAIN=W303;
RX MEDLINE=96004702; PubMed=7565597;
RA Kuessel P., Frasch M.;
RT "Yeast Srp1, a nuclear protein related to Drosophila and mouse
pendulin, is required for normal migration, division, and integrity
of nuclei during mitosis.";
RL Mol. Gen. Genet. 248:351-363(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA Obermaier B., Piravandi E., Rinke M., Domdey H.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP MUTAGENESIS.
RX MEDLINE=94316601; PubMed=8041713;
RA Yano R., Oakes M.L., Tabb M.M., Nomura M.;
RT "Yeast Srp1p has homology to armadillo/plakoglobin/beta-catenin and
participates in apparently multiple nuclear functions including the
maintenance of the nucleolar structure.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:6880-6884(1994).
RN [5]
RP IDENTIFICATION.
RX MEDLINE=95348058; PubMed=7622450;
RA Enekel C., Blobel G., Rexach M.;
RT "Identification of a yeast karyopherin heterodimer that targets
import substrate to mammalian nuclear pore complexes.";
RL J. Biol. Chem. 270:16499-16502(1995).
RN [6]
RP NUCLEOPORIN REPEAT BINDING REQUIREMENT.
RX MEDLINE=96097112; PubMed=8521485;
RA Rexach M., Blobel G.;
RT "Protein import into nuclei: association and dissociation reactions
involving transport substrate, transport factors, and nucleoporins.";
RL Cell 83:683-692(1995).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 89-510.
RX MEDLINE=98359119; PubMed=9695948;
RA Conti E., Uy M., Leighton L., Blobel G., Kuriyan J.;
RT "Crystallographic analysis of the recognition of a nuclear
localization signal by the nuclear import factor karyopherin alpha.";
RL Cell 94:193-204(1998).
CC -|- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
CC EITHER A SIMPLE OR BIPARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
CC SUBSTRATES TO THE NUCLEAR ENVELOPE. SEEMS TO ACT AS A CYTOSOLIC
CC RECEPTOR FOR BOTH SIMPLE AND BIPARTITE NLS MOTIFS (BY SIMILARITY).
CC -|- SUBUNIT: FORMS A COMPLEX WITH AN IMPORTIN BETA SUBUNIT. INTERACTS
CC WITH CSE1.
CC -|- SUBCELLULAR LOCATION: MAINLY LOCALIZED AT THE PERIPHERY OF THE
CC NUCLEUS.
CC -|- MISCELLANEOUS: BINDS TO NUCLEOPORIN FXFG BUT NOT GLFG REPEAT
CC REGIONS. RAN-GTP CAN DISRUPT THE KARYOPHERIN HETERODIMER BY
CC BINDING TO THE BETA SUBUNIT AND RELEASES BOTH SUBUNITS FROM THE
CC DOCKING SITE.

CC -|- SIMILARITY: BELONGS TO THE IMPORTIN ALPHA FAMILY.
CC -|- SIMILARITY: CONTAINS 8 ARM REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M75849; AAA35090.1; -;
DR EMBL; Z71465; CAA96083.1; -;
DR PIR; S30884; S30884.
DR PDB; 1BK5; 16-FEB-99.
DR PDB; 1BK6; 13-JAN-99.
DR SGD; S0005133; SRP1.
DR InterPro; IPR000225; Armadillo.
DR InterPro; IPR002652; Importina_B.
DR Pfam; PF00514; Armadillo_seg; 8.
DR Pfam; PF01749; IB; 1.
DR SMART; SM00185; ARM; 8.
DR PROSITE; PS50176; ARM_REPEAT; 2.
DR Transport; Protein transport; Repeat; Nuclear protein; 3D-structure.
FT DOMAIN 17 58 IB.
FT REPEAT 123 165 ARM 1.
FT REPEAT 166 207 ARM 2.
FT REPEAT 208 249 ARM 3.
FT REPEAT 250 291 ARM 4.
FT REPEAT 292 333 ARM 5.
FT REPEAT 334 375 ARM 6.
FT REPEAT 376 417 ARM 7.
FT REPEAT 418 462 ARM 8.
FT DOMAIN 463 542 ASP/GLU-RICH (ACIDIC).
FT MUTAGEN 116 116 S->F: IN SRP1-31; TS MUTANT; REDUCED
FT MUTAGEN 145 145 GROWTH RATE AND CHROMOSOME LOSS.
FT MUTAGEN 219 219 E->K: IN SRP1-49; TS MUTANT; ALTERATION
FT MUTAGEN 286 286 IN NUCLEOLAR AND MICROTUBULE MORPHOLOGY.
FT MUTAGEN 360 360 P->Q: IN SRP1-1; TS MUTANT.
FT MUTAGEN 459 459 D->N: IN SRP1-3; TS MUTANT.
FT SEQUENCE 542 AA; 60441 MM; 8D3A0CB76F2E7C00 CRC64; G->V: IN SRP1-54; TS MUTANT; REDUCED
GROWTH RATE.
SQ
Alignment Scores:
Pred. No.: 7.88 Length: 542
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 1 Gaps: 0
US-09-835-992A-20 (1-687) x IMA1_YEAST (1-542)
OY 497 TTACTTAGTTCTCCGAAGAAAC 520
Db 349 LeuLeuSerSerProLysGluAsn 356
RESULT 6
YS89_CAEEL STANDARD; PRT; 3178 AA.
AC Q09624; Q09625; Q969D4;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ZK945.9 in chromosome II.
GN ZK945.9/ZK945.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239; [1]

RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Wilkinson-Sproat J.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; 248544; CAB70192.1; -.
DR EMBL; 248582; CAB70192.1; JOINED.
DR EMBL; 248582; CAB70201.1; -.
DR EMBL; 248544; CAB70201.1; JOINED.
DR WormPep; 2K945.9; CE25697.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR001024; Lipoygenase_LH2.
DR InterPro; IPR000636; M-channel_nlg.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF01477; PLAT; 1.
DR Pfam; PF01825; GPS; 1.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00308; LH2; 1.
KW Hypothetical protein; Transmembrane.
FT DOMAIN 266 1196 SER/THR-RICH.
FT DOMAIN 1105 1241 GLY/SER-RICH.
FT DOMAIN 2071 2120 GPS.
FT TRANSMEM 13 30 POTENTIAL.
FT TRANSMEM 51 73 POTENTIAL.
FT TRANSMEM 2139 2161 POTENTIAL.
FT TRANSMEM 2348 2367 POTENTIAL.
FT TRANSMEM 2390 2412 POTENTIAL.
FT TRANSMEM 2451 2468 POTENTIAL.
FT TRANSMEM 2483 2505 POTENTIAL.
FT TRANSMEM 2567 2589 POTENTIAL.
FT TRANSMEM 2836 2858 POTENTIAL.
FT TRANSMEM 2939 2961 POTENTIAL.
FT TRANSMEM 2976 2998 POTENTIAL.
FT TRANSMEM 3038 3060 POTENTIAL.
SQ SEQUENCE 3178 AA; 344726 MW; F8239436D03666CD CRC64;

Alignment Scores:
Pred. No.: 5.61 Length: 3178
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-20 (1-687) x YS89_CAEEL (1-3178)
QY 398 AGATTTTAACTCNCNTAATTTTA 421
Db 2481 ArgPheLeuIleSerLeuIleLeu 2488

RESULT 7
YCF9_EUGGR STANDARD; PRT; 65 AA.
AC P32095;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 7.6 kDa protein ycf9 (ORF 65).

GN YCF9.
OS Euglena gracilis.
OG Chloroplast.
OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
OX NCBI_TaxID=3039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-2;
RX MEDLINE=93347989; PubMed=8346031;
RA Halliack R.B., Hong L., Drager R.G., Favreau M.R., Monfort A.,
RA Orsat B., Spielmann A., Stutz E.;
RT "Complete sequence of Euglena gracilis chloroplast DNA.";
RL Nucleic Acids Res. 21:3537-3544(1993).
CC -1- SIMILARITY: BELONGS TO THE YCF9 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; 211874; -, NOT_ANNOTATED_CDS.
DR EMBL; X70810; CAA50118.1; -.
DR PIR; S34538; S34538.
DR InterPro; IPR002644; Ycf9_struct.
DR Pfam; PF01737; YCF9; 1.
DR ProDom; PD004770; Ycf9_struct; 1.
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 65 AA; 7586 MW; 67DE2359D7A0F771 CRC64;

Alignment Scores:
Pred. No.: 120 Length: 65
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-20 (1-687) x YCF9_EUGGR (1-65)
QY 100 TTTTCTAGTTTCTAATA 80
Db 51 PhePheValPheLeuIle 57

RESULT 8
2SS1_ARATH STANDARD; PRT; 164 AA.
ID 2SS1_ARATH
AC P15457;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 2S seed storage protein 1 precursor (2S albumin storage protein)
DE (NMU2-2S albumin 1).
GN AT2S1 OR AT4G27140 OR T24A18.90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 38-73 AND 84-162.
RC STRAIN-cv. C24;
RA Krebbers E., Herdies L., de Clercq A., Seurinck J., Leemans J.,
RA van Damme J., Segura M., Gheysen G., van Montagu M.,
RA Vandekerckhove J.;
RT "Determination of the processing sites of an Arabidopsis 2S albumin
RT and characterization of the complete gene family.";
RL Plant Physiol. 87:859-866(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. C24;

RA Conceicao A.D.S., Krebbers E.;
 RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
 Harris B., Ansgore W., Brandt P., Grivell L.A., Rieger M.,
 Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
 Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 Vos P., Hohelsel J., Zimmermann W., Wedler H., Ridley P.,
 Langham S.-A., McCullagh B., Bilham L., Robben J.,
 Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
 Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 Weltzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
 Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
 Pettelt A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
 Neumann S., Argilou A., Vitale D., Liguori R., Piravandi E.,
 Massenot O., Quigley F., Clabaud G., Muendlein A., Feilber R.,
 Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 Chefor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 Perez-Perez A., Furnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C.,
 Antonou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 Chen E., Marra M., Martienssen R., McCombie W.R.;
 RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana.";
 RT Nature 402:769-777(1999).
 RL Nature 402:769-777(1999).
 CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
 CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 CC LINKED BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: THIS IS THE MOST ABUNDANT ISOFORM OF 2S ALBUMIN IN
 CC ARABIDOPSIS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M22032; AAA32743.1; -
 DR EMBL; Z24745; CAA80870.1; -
 DR EMBL; A13820; CAA01132.1; -
 DR EMBL; AL035680; CAB38844.1; -
 DR EMBL; AL161566; CAB79569.1; -
 DR PIR; JA0161; NWMU1.

DR PIR; S34676; S34676.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR000617; Napin.
 DR InterPro; IPR001768; Try/amy1_inhtr.
 DR Pfam; PF00234; tryp_alpha_aml; 1.
 DR PRINTS; PR00496; NAPIN.
 DR ProDom; PD002498; Napin; 1.
 DR SMART; SM00499; AAI; 1.
 KW Seed storage protein; Albumin; Signal; Multigene family.
 FT SIGNAL 1 21
 FT PROPEP 22 37
 FT CHAIN 38 73 2S SEED STORAGE PROTEIN 1 SMALL SUBUNIT.
 FT PROPEP 74 83
 FT CHAIN 84 162 2S SEED STORAGE PROTEIN 1 LARGE SUBUNIT.
 FT PROPEP 163 164
 SQ SEQUENCE 164 AA; 19014 MW; 2BF28CB474D9832B CRC64;
 Alignment Scores:
 Pred. No.: 101 Length: 164
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.30% Indels: 0
 DB: 1 Gaps: 0
 US-09-835-992a-20 (1-687) x 2SS1_ARATH (1-164)
 QY 482 TTGGCATGTGTTCTTACTT 502
 Db 12 LeuAlaLeuCySPhenLeuLeu 18
 |||||
 RESULT 9
 2SS3_ARATH STANDARD; PRT; 164 AA.
 ID 2SS3_ARATH
 AC P15459;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 2S seed storage protein 3 precursor (2S albumin storage protein)
 DE (NWMU2-2S albumin 3).
 GN AT2S3 OR AT4G27160 OR T24A18.110.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. C24;
 RA krebbes E., Herdies L., de Clercq A., Seurinck J., Leemans J.,
 van Damme J., Segura M., Gheysen G., van Montagu M.,
 Vandekerckhove J.;
 RT "Determination of the processing sites of an Arabidopsis 2S albumin
 RT and characterization of the complete gene family.";
 RL Plant Physiol. 87:859-866(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. C24;
 RA Conceicao A.D.S., Krebbers E.;
 RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
 Harris B., Ansgore W., Brandt P., Grivell L.A., Rieger M.,
 Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
 Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 Vos P., Hohelsel J., Zimmermann W., Wedler H., Ridley P.,
 Langham S.-A., McCullagh B., Bilham L., Robben J.,
 Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,

RA Braeken M., Weltjens I., Voet M., Bastlaens I., Aert R., Defoor E.,
RA Weltzenegger T., Bothe G., Ramsperger U., Hildert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Mooljman P., Klein Lankhorst R., Rose M., Hauf J., Koelter P.,
RA Bernolser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buysaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
RA Petteit A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloecker H., Scherfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Grandrath K., Dauner D., Herzl A.,
RA Neumann S., Argirliou A., Vltale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefod F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Glibons T., Weber N., Vandebol M., Barges M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bieleke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Mlnx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Gelsel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C.,
RA Antonolu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lochl M., Johnson A.,
RA Chen E., Marra M., Martensen R., McComble W.R.,
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana."
RL Nature 402:769-777(1999).
RN [4]
RP SEQUENCE OF 103-164 FROM N.A.
RC STRAIN-cv. Columbia;
RA Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.;
RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
CC LINKED BY DISULFIDE BONDS.
CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M22035; AAA32745.1; -
DR EMBL: Z24744; CAA80868.1; -
DR EMBL: AL035680; CAB38846.1; -
DR EMBL: AL161566; CAB79571.1; -
DR EMBL: Z17580; CAA79001.1; -
DR PIR: JAO163; NMMU3.
DR PIR: S34674; S34674.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR000617; Napin.
DR InterPro: IPR001768; Try/amy1_inhbt.
DR Pfam: PF00234; try_alpha_aml; 1.
DR PRINTS: PR00496; NAPIN.
DR PRODOM: PD002498; Napin; 1.
DR SMART: SM00499; AAI; 1.
KW Seed storage protein; Albumin; Signal; Multigene family.
FT SIGNAL 1 21
FT PROPEP 22 37
FT CHAIN 38 72 2S SEED STORAGE PROTEIN 3 SMALL SUBUNIT

FT PROPEP 73 81 (BY SIMILARITY).
FT CHAIN 82 164
FT SEQUENCE 164 AA; 18762 MW; C9BEB6718549F248 CRC64;
SQ
Alignment Scores:
Pred. No.: 101 Length: 164
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 1 Gaps: 0
US-09-835-992a-20 (1-687) x 2SS3_ARATH (1-164)
OY 482 TTGCGATGCTTCTACTT 502
DB 12 LeuAlaLeucysphLeuLeu 18
RESULT 10
YPKW_THECU
ID YPKW_THECU STANDARD; PRT; 180 AA.
AC P49694;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein in pkwa 5'region (ORF1) (Fragment).
OS Thermomonospora curvata.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptosporangineae; Thermomonosporaceae;
OC Thermomonospora.
OX NCBI_TaxID=2020;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CCM 3352;
RX MEDLINE=96200125; PubMed=8631732;
RA Janda L., Tichy P., Spizek J., Petricek M.;
RT "A deduced Thermomonospora curvata protein containing
RT serine/threonine protein kinase and WD-repeat domains."
RL J. Bacteriol. 178:1487-1489(1996).
CC -1- SIMILARITY: BELONGS TO THE CDAR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF115313; -; NOT_ANNOTATED_CDS.
DR InterPro: IPR000712; Bcl2_BH.
KW Hypothetical protein.
FT NON_TER 1 1
FT SEQUENCE 180 AA; 19880 MW; 7399D55F9E204D40 CRC64;
SQ
Alignment Scores:
Pred. No.: 98.8 Length: 180
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 1 Gaps: 0
US-09-835-992a-20 (1-687) x YPKW_THECU (1-180)
OY 62 AATTCAGTGAGATATAGATA 42
DB 132 AsnSerValArgTyrArgIle 138
RESULT 11
YPBH_BACSU


```

ID YPBH_BACSU STANDARD; PRT; 194 AA.
AC P50734;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ypbH.
GN ypbH.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serror P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
RT the sera and kdg loci cloned in a yeast artificial chromosome.";
RL Microbiology 142:2005-2016(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone C.,
RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambut R., Wedler E., Wedler H., Weltzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L47648; AAC83952.1; -
DR EMBL; Z99115; CAB14213.1; -
DR EMBL; Z99116; CAB14229.1; -
DR Subtilist; BG11434; ypbH.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 194 AA; 22157 MW; B27C6ABE4EA0745C CRC64;

```

```

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.32%
DB: 1
Gaps: 0
US-09-835-992A-20 (1-687) x YPBH_BACSU (1-194)
QY 303 AAATATAATATTTTAAACA 283
Db 10 LysileLysilepheLeuThr 16
RESULT 12
YAHF_ECOLI STANDARD; PRT; 201 AA.
ID YAHF_ECOLI
AC P77736;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical ANK-repeats protein yahd.
GN YAHF OR B0318.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000139; AAC73421.1; -
DR EMBL; U73857; AAB18044.1; -
DR EcoGene; EG13588; yahd.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 6.
DR SMART; SM00248; ANK; 2.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Hypothetical protein; Repeat; ANK repeat; Complete proteome.
FT REPEAT 5 34 ANK 1.
FT REPEAT 38 67 ANK 2.
FT REPEAT 71 100 ANK 3.
FT REPEAT 104 134 ANK 4.
FT REPEAT 138 172 ANK 5.
FT REPEAT 176 201 ANK 6.
SQ SEQUENCE 201 AA; 21687 MW; A90244CEB892E8BD CRC64;

```

Alignment Scores:

Pred. No.:	97.4	Length:	194
Score:	7.00	Matches:	7

Alignment Scores:

Pred. No.:	96.7	Length:	201
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0

DB: 1 Gaps: 0
US-09-835-992A-20 (1-687) x YAHU_ECOLI (1-201)
QY 312 GCCTGTGTCAAGCTCTGATC 332
|||||
Db 53 AlAcysValGlnAlaLeuIle 59
RESULT 13
FCEL_RAT STANDARD; PRT; 245 AA.
AC P12371;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity Immunoglobulin epsilon receptor alpha-subunit precursor
DE (FCEI) (IGE FC receptor, alpha-subunit) (FC-epsilon RI-alpha).
GN FCEIA OR FCEIA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid-10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88024987; PubMed-2959318;
RT Kinet J.-P., Metzger H., Hakim J., Kochan J.;
RT "A cDNA presumptively coding for the alpha subunit of the receptor
RT with high affinity for immunoglobulin E.";
RL Biochemistry 26:4605-4610(1987).
RN [2]
RP REVISIONS.
RA Kochan J.;
RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-Mast cells;
RX MEDLINE-88158102; PubMed-2964640;
RA Shimizu A., Tepler I., Benfey P.N., Berenstein E.H., Siraganian R.P.,
RA Leder P.;
RT "Human and rat mast cell high-affinity immunoglobulin E receptors:
RT characterization of putative alpha-chain gene products.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1907-1911(1988).
RN [4]
RP SEQUENCE OF 21-245 FROM N.A.
RX MEDLINE-88289772; PubMed-2969594;
RA Liu F.-T., Albrandt K., Robertson M.W.;
RT "cDNA heterogeneity suggests structural variants related to the high-
RT affinity IGE receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:5639-5643(1988).
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
CC DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17153; AAA42045.1; -
DR EMBL; J03606; AAA41582.1; -
DR EMBL; M21622; AAA41146.1; -
DR PIR; A27116; A27116.
DR PIR; A30154; A30154.

DR PTR; C31327; C31327.
DR HSSP; P12319; IALS.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; Ig; 2.
KW IGE-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1 23
FT CHAIN 24 245
FT FT
FT DOMAIN 24 204
FT TRANSMEM 205 223
FT DOMAIN 224 245
FT DOMAIN 42 98
FT DOMAIN 123 181
FT DISULFID 49 91
FT DISULFID 130 174
FT CARBOHYD 52 52
FT CARBOHYD 53 53
FT CARBOHYD 58 58
FT CARBOHYD 65 65
FT CARBOHYD 123 123
FT CARBOHYD 158 158
FT CARBOHYD 167 167
SQ SEQUENCE 245 AA; 27793 MW; A0E67DD363B72197 CRC64;
Alignment Scores:
Pred. No.: 93.1 Length: 245
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 1 Gaps: 0
US-09-835-992A-20 (1-687) x FCEL_RAT (1-245)
QY 69 AATAACATTAGTATTAGAA 89
|||||
Db 157 AsnAsnIleSerIleArgLys 163
RESULT 14
PYRF_PICST
ID PYRF_PICST STANDARD; PRT; 267 AA.
AC P49434;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
DE (OMPDCase) (OMPdecase) (Uridine 5'-monophosphate synthase) (UMP
DE synthase).
GN URA3.
OS Pichia stipitis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_Taxid-4924;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 58785 / CBS 6054;
RX MEDLINE-95110115; PubMed-7811063;
RA Yang V.W., Marks J.A., Davis B.P., Jeffries T.W.;
RT "High-efficiency transformation of Pichia stipitis based on its URA3
RT gene and a homologous autonomous replication sequence, ARS2.";
RL Appl. Environ. Microbiol. 60:4245-4254(1994).
CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate - UMP + CO(2).
CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U08629; AAA65978.1; -.
DR HSSP; P03962; 1DQW.
DR InterPro; IPR001754; OMPdecase.
DR Pfam; PF00215; OMPdecase; 1.
DR PROSITE; PS00156; OMPDECASE; 1.
KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
FT ACT_SITE 94 94 BY SIMILARITY.
SQ SEQUENCE 267 AA; 29529 MW; D98361C398F49E06 CRC64;

Alignment Scores:

Pred. No.:	91.6	Length:	267
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	1	Gaps:	0

US-09-835-992A-20 (1-687) x PYRF_PICST (1-267)

QY 507 GAACTAAGTAGAACAACAAT 487

Db 80 GluLeuSerLysLysHisAsn 86

RESULT 15

PYRF_CANTR
ID PYRF_CANTR STANDARD; PRT; 268 AA.
AC 042771;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)

DE (OMPDCase) (OMPdecase) (Uridine 5'-monophosphate synthase) (UMP

GN URA3.

OS Candida tropicalis (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; mitosporic Saccharomycetales; Candida.

OX NCBL_TaxID=5482;

RL Curr. Microbiol. 37:210-213(1998).

CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate - UMP + CO(2).

CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.

CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; AF040702; AAB96773.1; -.
DR HSSP; P03962; 1DQW.
DR InterPro; IPR001754; OMPdecase.
DR Pfam; PF00215; OMPdecase; 1.
DR PROSITE; PS00156; OMPDECASE; 1.
KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
FT ACT_SITE 94 94 BY SIMILARITY.
SQ SEQUENCE 268 AA; 29672 MW; A43CC1ED1CB24237 CRC64;

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	1	Gaps:	0

US-09-835-992A-20 (1-687) x PYRF_CANTR (1-268)

QY 507 GAACTAAGTAGAACAACAAT 487

Db 80 GluLeuSerLysLysHisAsn 86

Search completed: January 14, 2003, 17:29:11
Job time : 14.6084 secs

Alignment Scores:

Pred. No.:	91.5	Length:	268
Score:	7.00	Matches:	7

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:20:59 ; Search time 48.3743 Seconds
(without alignments)
5852.461 Million cell updates/sec

Title: US-09-835-992A-20

Perfect score: 212

Sequence: 1 attttaatttctataaac.....cttcgtatgcacagtgtt 687

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlp
-O=/cgn2_1/USPRO_spool/US09835992/runat_14012003_161526_4027/app_query.fasta_1.3932
-DB=SPTREMBL_21 -QFMT=fastan -SUFFIX=olg.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdl
-LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992 @CGN_1_1_346 @runat_14012003_161526_4027 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	8	3.8	151	6	O97662
					097662 ovls arles

C	2	8	3.8	280	5	O9N5G4	O9N5G4 caenorhabd1
C	3	8	3.8	342	6	O9GK76	O9GK76 capra hircu
C	4	8	3.8	342	6	O9TTY5	O9TTY5 bos taurus
C	5	8	3.8	343	5	O17820	O17820 caenorhabd1
C	6	8	3.8	352	17	O8TTZ5	O8TTZ5 methanosarc
	7	8	3.8	376	16	O92RV3	O92RV3 rhizobium m
	8	8	3.8	512	16	O9HVA7	O9HVA7 pseudomonas
	9	8	3.8	791	17	O9YB89	O9YB89 aeropyrum p
C	10	8	3.8	1183	16	O9BQH9	O9BQH9 mycoplasma
C	11	8	3.8	3178	5	O969D4	O969D4 caenorhabd1
C	12	7	3.3	55	10	O41910	O41910 arabidopsis
C	13	7	3.3	55	12	O91FK1	O91FK1 chilo iride
C	14	7	3.3	55	16	O92UP8	O92UP8 rhizobium m
C	15	7	3.3	65	11	O9D860	O9D860 mus musculu
C	16	7	3.3	65	11	O9CQP0	O9CQP0 mus musculu
C	17	7	3.3	72	6	O9GMR4	O9GMR4 macaca fasc
C	18	7	3.3	85	2	O9FCV5	O9FCV5 lactobacill1
C	19	7	3.3	85	5	O9VDM5	O9VDM5 drosophila
C	20	7	3.3	91	8	O99813	O99813 lxodes hexa
C	21	7	3.3	105	16	O9K7E6	O9K7E6 bacillus ha
C	22	7	3.3	106	2	O47966	O47966 herpetosiph
C	23	7	3.3	117	8	O09313	O09313 nesticus st
C	24	7	3.3	117	8	O09314	O09314 nesticus st
C	25	7	3.3	117	8	O09317	O09317 nesticus ba
C	26	7	3.3	117	8	O09333	O09333 nesticus sl
C	27	7	3.3	120	12	O57169	O57169 vaccinia vl
C	28	7	3.3	121	2	O9AL17	O9AL17 enterococcu
C	29	7	3.3	123	17	O973P2	O973P2 sulfolobus
C	30	7	3.3	123	17	O96ZW9	O96ZW9 sulfolobus
C	31	7	3.3	124	12	P87608	P87608 cowpox viru
C	32	7	3.3	126	12	O89067	O89067 variola vir
C	33	7	3.3	172	5	O95XJ8	O95XJ8 caenorhabd1
C	34	7	3.3	183	10	O8RWP7	O8RWP7 arabidopsis
C	35	7	3.3	191	12	O9DVU7	O9DVU7 plutella xy
C	36	7	3.3	195	17	O59004	O59004 pyrococcus
C	37	7	3.3	198	6	O9GKZ3	O9GKZ3 sus scrofa
C	38	7	3.3	201	16	O8X6B2	O8X6B2 escherichia
C	39	7	3.3	205	10	O9SVW1	O9SVW1 arabidopsis
C	40	7	3.3	207	5	O22044	O22044 caenorhabd1
C	41	7	3.3	208	11	O91ZW7	O91ZW7 mus musculu
C	42	7	3.3	225	12	O8VZJ1	O8VZJ1 camelipox vl
C	43	7	3.3	228	5	O9NSD7	O9NSD7 caenorhabd1
C	44	7	3.3	236	10	O9AW49	O9AW49 guillardi
C	45	7	3.3	238	2	O8VTA2	O8VTA2 agrobacteri1

ALIGNMENTS

RESULT 1							
O97662		PRELIMINARY;	PRT;	151	AA.		
ID	O97662						
AC	O97662;						
DT	01-MAY-1999 (TREMBlrel. 10, Created)						
DT	01-MAY-1999 (TREMBlrel. 10, Last sequence update)						
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)						
DE	Platelet activating factor receptor (Fragment).						
OS	Ovis aries (Sheep).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;						
OC	Bovidae; Caprinae; Ovis.						
OX	NCBI_Taxid=9940;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	TISSUE=LUNG;						
RA	Sander F.C., Ibe O.B., Raj J.U.;						
RT	"Developmentally Regulated Expression of Platelet Activating Factor						
RT	Receptor in Ovine Lung."						
RL	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; AF099674; AAC77459.1; -						
DR	InterPro; IPR000276; GPCR_Rhodpsn.						
DR	Pfam; PF00001; 7tm_1; 1.						
DR	PRINTS; PR00237; GPCR_RHODOPSN.						
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.						

DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 151 151
SQ SEQUENCE 151 AA; 17001 MW; 31ECAFA2006AF951C CRC64;

Alignment Scores:

Pred. No.:	30.1	Length:	151
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.79%	Indels:	0
DB:	6	Gaps:	0

US-09-835-992A-20 (1-687) x O97662 (1-151)

OY 103 TCTTTTTCCTACTTTTCTAATA 80
|||||
DB 138 SerpPheLeuValPheLeuIle 145

RESULT 2

O9N5G4 PRELIMINARY; PRT; 280 AA.

AC O9N5G4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 29.9 kDa protein.
GN R05D8.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Langston Y., Wohlmann P., Lennox S.;
RT "The sequence of C. elegans cosmid R05D8.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
(SDR) FAMILY.
CC EMBL; AC006677; AAF39950.1; -.
DR HSSP; P50162; IAE1.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Hypothetical protein; Oxidoreductase.
SQ SEQUENCE 280 AA; 29923 MW; 588C9AE6F1D5F006 CRC64;

Alignment Scores:

Pred. No.:	26.6	Length:	280
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	5	Gaps:	0

US-09-835-992A-20 (1-687) x O9N5G4 (1-280)

OY 516 AAACTCTTAAGTAATCTTCAG 539
|||||
DB 116 LysThrLeuLysLeuAsnLeuGln 123

RESULT 3

O9GK76 PRELIMINARY; PRT; 342 AA.

AC O9GK76;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Platelet-activating factor receptor.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang W., Diehl J.R., Piumi F.;
RT "Partial Sequence of Caprine Platelet-Activating Factor Receptor
Gene.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF302764; AAG39982.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR01559; DUFFYANTIGEN.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 342 AA; 39612 MW; 39D9379AED60FD4E CRC64;

Alignment Scores:

Pred. No.:	25.6	Length:	342
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.79%	Indels:	0
DB:	6	Gaps:	0

US-09-835-992A-20 (1-687) x O9GK76 (1-342)

OY 103 TCTTTTTCCTACTTTTCTAATA 80
|||||
DB 194 SerpPheLeuValPheLeuIle 201

RESULT 4

O9TTY5 PRELIMINARY; PRT; 342 AA.

AC O9TTY5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Platelet-activating factor receptor.
GN PAFR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang W.S., Diehl J.R., Roubush W.E.;
RT "Partial Sequence of Bovine Platelet-Activating Factor Receptor
Gene.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA He B., Tiemann U., Kanitz W., Weikard R., Laurent P., Schwerin M.,
Schmidt P.;
RT "Molecular characterization of bovine platelet-activating factor


```
RT receptor transcripts and their detection in different tissues of
RT catle.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF187321; AAF01439.2; -.
DR EMBL: AJ295321; CAC43290.1; -.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PRO1559; DUFFYANTIGEN.
DR PRINTS: PRO0237; GPCRRIHODPSN.
DR PROSITE: PS00237; G_PROTEIN_RECCEP_F1_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECCEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 342 AA; 39691 MW; 7C4236205AE937C9 CRC64;

Alignment Scores:
Pred. No.: 25.6 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: Gaps: 0

US-09-835-992A-20 (1-687) x Q9TTY5 (1-342)
OY 103 TCTTTTCTCTAGTTTCTAATA 80
Db 194 SerpHeLeuValPheLeuile 201

RESULT 5
Q17820 PRELIMINARY; PRT; 343 AA.
AC Q17820;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE C08B6.9 protein.
GN C08B6.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilkinson J.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z72502; CAA96591.1; -.
DR InterPro: IPR000594; ThIF_domain.
DR Pfam: PF00899; ThIF; 1.
SQ SEQUENCE 343 AA; 38738 MW; 2897C0F918E81B6 CRC64;

Alignment Scores:
Pred. No.: 25.6 Length: 343
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: Gaps: 0

US-09-835-992A-20 (1-687) x Q17820 (1-343)
OY 628 CACAATAAATTAGGACACTAAA 605
Db 23 GlnAsnLysIleArgAsnSerLys 30

RESULT 6
Q8TTZ5 PRELIMINARY; PRT; 352 AA.
AC Q8TTZ5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sulfate/molybdate ABC transporter, solute-binding protein.
GN MA0280.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Alnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArelano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kueltner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL: AE010687; AAM03733.1; -.
KW Complete proteome.
SQ SEQUENCE 352 AA; 38715 MW; 15652724987E8A9E CRC64;

Alignment Scores:
Pred. No.: 25.4 Length: 352
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: Gaps: 0

US-09-835-992A-20 (1-687) x Q8TTZ5 (1-352)
OY 100 TTTTCTAGTTTCTAATACTA 77
Db 10 PhePheLeuValPheLeuileLeu 17

RESULT 7
Q92RV3 PRELIMINARY; PRT; 376 AA.
AC Q92RV3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative permease protein.
GN R00738 OR SMC00793.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetlelle D., Puhler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL: AL591784; CAC45310.1; -.

```

DR InterPro; IPR000560; HisAc_phsphtse.
DR InterPro; IPR002549; UPF0118.
DR Pfam; PF01594; UPF0118; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 376 AA; 40673 MW; F119EE2DFB8CE87 CRC64;

Alignment Scores:
Pred. No.: 25.1 Length: 376
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 16 Gaps: 0

US-09-835-992A-20 (1-687) x Q9HVA7 (1-376)

OY 479 CTGTTGGCATTGTTCTTACTT 502
Db 97 LeuLeuAlaLeuCysPheLeu 104

RESULT 8
Q9HVA7 PRELIMINARY; PRT; 512 AA.

ID Q9HVA7;
AC Q9HVA7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Iron (III)-transport system permease HltB.
GN HITB OR PA4688.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -1- FUNCTION: PROBABLY PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT
CC SYSTEM. PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE
CC SUBSTRATE ACROSS THE MEMBRANE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS.
DR EMBL; AE004882; AAG08075.1; -.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp. 2.
DR PROSITE; PS00402; BPD_TRANS_P_NN_MEMBR; 1.
KW Transmembrane; Transport; Complete proteome.
SQ SEQUENCE 512 AA; 56176 MW; 9A21A90B3CDFBF0A CRC64;

Alignment Scores:
Pred. No.: 23.6 Length: 512
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 16 Gaps: 0

US-09-835-992A-20 (1-687) x Q9HVA7 (1-512)

OY 479 CTGTTGGCATTGTTCTTACTT 502
Db 230 LeuLeuAlaLeuCysPheLeu 237

RESULT 9
Q9YB89 PRELIMINARY; PRT; 791 AA.

ID Q9YB89;
AC Q9YB89;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein APE1708.
GN APE1708.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcates;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1."
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000062; BAA80709.1; -.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR002570; Hpt.
DR SMART; SM00073; HPT; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 791 AA; 85044 MW; D52098F863EBA82F CRC64;

Alignment Scores:
Pred. No.: 21.6 Length: 791
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 17 Gaps: 0

US-09-835-992A-20 (1-687) x Q9YB89 (1-791)

OY 76 TTAGTATTAGAAAAGTAGGAAA 99
Db 134 LeuValLeuGluLysLeuGlyLys 141

RESULT 10
Q98QH9 PRELIMINARY; PRT; 1183 AA.

ID Q98QH9;
AC Q98QH9;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Hypothetical protein MYPu_3820.
GN MYPu_3820.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis."
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL; AL445564; CAC13555.1; -.
DR Mypulist; MYPu_3820; -.

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1183 AA; 140128 MW; 565C3A6948E1A9D0 CRC64;
Alignment Scores:
Pred. No.: 20 Length: 1183
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 16 Gaps: 0
US-09-835-992A-20 (1-687) x Q96QH9 (1-1183)
QY 194 TTTTCAATATATATCTCATTTCA 171
Db 577 PheSerAsnTyrAsnLeuIleSer 584
RESULT 11
Q969D4 PRELIMINARY; PRT; 3178 AA.
AC Q969D4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Lov-1 protein.
GN LOV-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilkison J.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; 248582; CAB70201.1; -.
DR EMBL; 248544; CAB70201.1; JOINED.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR001024; Lipoxigenase_LH2.
DR InterPro; IPR000636; M-channel_nlg.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF01477; PLAT; 1.
SQ SEQUENCE 3178 AA; 344726 MW; F8239436D03666CD CRC64;
Alignment Scores:
Pred. No.: 16.4 Length: 3178
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 5 Gaps: 0
US-09-835-992A-20 (1-687) x Q969D4 (1-3178)
QY 398 AGATTTTAACTCTCNCCTAATTTTA 421
Db 2481 ArgPheLeuIleSerLeuIleLeu 2488
RESULT 12
Q41910 PRELIMINARY; PRT; 55 AA.
ID Q41910;
AC Q41910;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE 2S seed storage protein 3 (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-GREEN SILICONES OF A. THALIANA ECOTYPE COLUMBIA;
RA Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.;
RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z17640; CAA79021.1; -.
FT NON_TER 55
SQ SEQUENCE 55 AA; 6248 MW; F4B7740140803E95 CRC64;
Alignment Scores:
Pred. No.: 411 Length: 55
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 10 Gaps: 0
US-09-835-992A-20 (1-687) x Q41910 (1-55)
QY 482 TTGGCATGTGTTCTTACTT 502
Db 12 LeuAlaLeuCysPheLeuLeu 18
RESULT 13
Q91FK1
ID Q91FK1 PRELIMINARY; PRT; 55 AA.
AC Q91FK1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 323L.
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=10488;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Darai G., Fluegel R.M.;
RT "DNA analysis of insect iridescent virus 6: evidence for circular
RT permutation and terminal redundancy."
RL J. Virol. 49:609-614(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86174607; PubMed=3959991;
RA Lorbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;
RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis
RT in mice."
RL Med. Microbiol. Immunol. 175:43-53(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87321126; PubMed=2820141;
RA Schmitzler P., Soltan J.B., Fischer M., Relsner H., Scholz J.,
RA Delius H., Darai G.;
RT "Molecular cloning and physical mapping of the genome of insect
RT iridescent virus type 6: further evidence for circular permutation of
RT the viral genome."
RL Virology 160:66-74(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89073752; PubMed=3201750;
RA Fischer M., Schmitzler P., Delius H., Darai G.;
RT "Identification and characterization of the repetitive DNA element in
RT the genome of insect iridescent virus type 6."
RL Virology 167:485-496(1988).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92196996; PubMed=1549908;
RA Handermann M., Schmitzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,

RA Darai G.;
RT "Identification and mapping of origins of DNA replication within the
RT DNA sequences of the genome of insect iridescent virus type 6.";
RL Virus Genes 6:19-32(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93118242; PubMed=1475907;
RA Sonntag K.C., Darai G.;
RT "Characterization of the third origin of DNA replication of the genome
RT of insect iridescent virus type 6.";
RL Virus Genes 6:333-342(1992).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=93260401; PubMed=8492091;
RA Stohwasser R., Raab K., Schmitzler P., Janssen W., Darai G.;
RT "Identification of the gene encoding the major capsid protein of
RT insect iridescent virus type 6 by polymerase chain reaction.";
RL J. Gen. Virol. 74:873-879(1993).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=94167241; PubMed=8121799;
RA Schmitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
RA Delius H., Darai G.;
RT "Identification of genes encoding zinc finger proteins, non-histone
RT chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
RT in the genome of Chilo iridescent virus.";
RL Nucleic Acids Res. 22:158-166(1994).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=94353641; PubMed=8073636;
RA Sonntag K.C., Schmitzler P., Koonin E.V., Darai G.;
RT "Chilo iridescent virus encodes a putative helicase belonging to a
RT distinct family within the 'DEAD/H' superfamily: implications for the
RT evolution of large DNA viruses.";
RL Virus Genes 8:151-158(1994).
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=95213160; PubMed=7698884;
RA Sonntag K.C., Schmitzler P., Janssen W., Darai G.;
RT "Identification of the primary structure and the coding capacity of
RT the genome of insect iridescent virus type 6 between the genome
RT coordinates 0.310 and 0.347 (7990 bp).";
RL Intervirology 37:287-297(1994).
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=94292906; PubMed=8021587;
RA Schmitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
RA Koonin E.V., Darai G.;
RT "Insect iridescent virus type 6 encodes a polypeptide related to the
RT largest subunit of eukaryotic RNA polymerase II.";
RL J. Gen. Virol. 75:1557-1567(1994).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=98141693; PubMed=9482589;
RA Bahr U., Tidona C.A., Darai G.;
RT "The DNA sequence of Chilo iridescent virus between the genome
RT coordinates 0.101 and 0.391; similarities in coding strategy between
RT insect and vertebrate iridoviruses.";
RL Virus Genes 15:235-245(1997).
RN [13]
RP SEQUENCE FROM N.A.
RX MEDLINE=99125223; PubMed=9926400;
RA Muller K., Tidona C.A., Bahr U., Darai G.;
RT "Identification of a thymidylate synthase gene within the genome of
RT Chilo iridescent virus.";
RL Virus Genes 17:243-258(1998).
RN [14]
RP SEQUENCE FROM N.A.
RX MEDLINE=99383793; PubMed=10456793;
RA Muller K., Tidona C.A., Darai G.;
RT "Identification of a gene cluster within the genome of Chilo
RT iridescent virus encoding enzymes involved in viral DNA replication
RT and processing.";

RL Virus Genes 18:243-264(1999).
RN [15]
RP SEQUENCE FROM N.A.
RX MEDLINE=21342589; PubMed=11448171;
RA Jakob N.J., Muller K., Bahr U., Darai G.;
RT "Analysis of the First Complete DNA Sequence of an Invertebrate
RT Iridovirus: Coding Strategy of the Genome of Chilo Irdescent Virus.";
RL Virology 286:182-196(2001).
RN [16]
RP SEQUENCE FROM N.A.
RA Jakob N.J., Mueller K., Bahr U., Darai G.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF303741; AAK82184.1;
SQ SEQUENCE 55 AA; 6886 MW; 100D465DFB142CAE CRC64;

Alignment Scores:
Pred. No.: 411 Length: 55
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 12 Gaps: 0

US-09-835-992A-20 (1-687) x Q91FK1 (1-55)
QY 602 CCAAGAGACTATGTAAACAA 582
Db 13 ProlysAsnTyValAsnLys 19

RESULT 14
Q92UP8 PRELIMINARY; PRT; 55 AA.
ID Q92UP8
AC Q92UP8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein RB1047.
GN RB1047 OR SMB21672.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weldner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoeelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puenhler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL: AL603645; CAC49447.1;
KW plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 55 AA; 6547 MW; E20E5EE35361C1CB CRC64;

Alignment Scores:
Pred. No.: 411 Length: 55
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 16 Gaps: 0

US-09-835-992A-20 (1-687) x Q92UP8 (1-55)
QY 529 AGTTAAGAGTTTCTCTGGG 509
Db 25 SerLeuArgValPheLeuGly 31

RESULT 15
Q9DB60

ID	Q9DB60	PRELIMINARY;	PRT;	65 AA.
AC	Q9DB60;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)			
DE	0610009M10R1k protein.			
	GN MRP133 OR 0610009M10R1K.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID-10090;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-C57BL/6J; TISSUE=SMALL INTESTINE;			
RX	MEDLINE-21085660; PubMed-11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	AbeKawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Salto R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Glissl C., Kling B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,			
RA	Hayashizaki Y.;			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RL	Nature 409:685-690(2001).			
DR	EMBL; AK008431; BAB25655.1; -.			
DR	MGI; MGI:2137225; MRP133.			
SO	SEQUENCE 65 AA; 7344 MW; 358E69CA0359B6C CRC64;			

Alignment Scores:	
Pred. No.:	397
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	3.32%
DB:	11
Length:	65
Matches:	7
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-835-992A-20 (1-687) x Q9D860 (1-65)

QY 590 GTAAACAAAAGTTTGT TT 570

Db 50 ValAsnLysValLeuPhe 56

Search completed: January 14, 2003, 17:36:55
Job time : 54.3743 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:24 ; Search time 48.9632 Seconds
(without alignments)
5410.226 Million cell updates/sec

Title: US-09-835-992A-21
Perfect score: 314
Sequence: 1 ctccaccagtgtctctctcag.....ccaccctgggtgtgtgat 994

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:
-MODEL=frame+_{n2p}.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09835992/runat_14012003_161525_3979/app_query.fasta_1.3932
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=olig.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdl
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992.@cgn_1_1_121.@runat_14012003_161525_3979 -NCPD=6 -ICPD=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	15.9	547	23 ABB57301	Mouse ischaemic co
2	30	9.6	378	22 ABB65147	Drosophila melanog
3	30	9.6	544	22 ABB65056	Drosophila melanog
4	9	2.9	397	22 AAE09042	Equine influenza v
5	9	2.9	716	22 AAE09044	Equine influenza v
6	8	2.5	153	21 AAG20138	Arabidopsis thalia
7	8	2.5	182	21 AAG39091	Arabidopsis thalia
8	8	2.5	187	21 AAG20137	Arabidopsis thalia
9	8	2.5	235	21 AAG20136	Arabidopsis thalia
10	8	2.5	241	21 AAG37797	Arabidopsis thalia
11	8	2.5	248	21 AAG39090	Arabidopsis thalia
12	8	2.5	261	21 AAG37796	Arabidopsis thalia
13	8	2.5	284	21 AAG39054	Arabidopsis thalia
14	8	2.5	291	21 AAG11127	Arabidopsis thalia
15	8	2.5	291	21 AAG37795	Arabidopsis thalia
16	8	2.5	295	21 AAG30200	Arabidopsis thalia
17	8	2.5	295	21 AAG44684	Arabidopsis thalia
18	8	2.5	301	21 AAG39089	Zea mays protein f
19	8	2.6	302	21 AAG22946	Arabidopsis thalia
20	8	2.6	306	21 AAG22945	Arabidopsis thalia
21	8	2.6	314	21 AAG22944	Arabidopsis thalia
22	8	2.6	314	21 AAG39926	Arabidopsis thalia
23	8	2.6	314	23 ABB93213	Herbicidally activ
24	8	2.5	317	21 AAG36745	Arabidopsis thalia
25	8	2.5	325	21 AAG39925	Arabidopsis thalia
26	8	2.6	330	21 AAG39053	Arabidopsis thalia
27	8	2.6	333	21 AAG39924	Arabidopsis thalia
28	8	2.5	334	21 AAG14646	Arabidopsis thalia
29	8	2.5	345	21 AAG11126	Arabidopsis thalia
30	8	2.5	345	21 AAG30199	Arabidopsis thalia
31	8	2.5	345	21 AAG44683	Arabidopsis thalia
32	8	2.5	383	21 AAG36744	Arabidopsis thalia
33	8	2.5	384	13 AAR24296	Regulatory protein
34	8	2.5	397	22 AAU36868	Staphylococcus aur
35	8	2.5	411	21 AAG39052	Arabidopsis thalia
36	8	2.5	414	21 AAG11125	Arabidopsis thalia
37	8	2.5	414	21 AAG39458	Arabidopsis thalia
38	8	2.5	427	22 AAB20010	Brassica 3-ketocoy
39	8	2.5	436	21 AAG36743	Arabidopsis thalia
40	8	2.5	436	23 ABB92008	Herbicidally activ
41	8	2.5	443	21 AAG14645	Arabidopsis thalia
42	8	2.5	443	22 AAB20005	Arabidopsis 3-keto
43	8	2.5	446	21 AAG14644	Arabidopsis thalia
44	8	2.5	453	21 AAG39457	Arabidopsis thalia
45	8	2.5	457	21 AAG39456	Arabidopsis thalia

ALIGNMENTS

RESULT 1
/ABB57301
ID ABB57301 standard; Protein: 547 AA.
AC ABB57301;
XX DT 07-MAR-2002 (first entry)
XX DE Mouse ischaemic condition related protein sequence SEQ ID NO:843.
XX DE Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX OS Mus musculus.
XX PN WO200188188-A2.
XX PD 22-NOV-2001.
XX

18-MAY-2001; 2001WO-JP04192.
18-MAY-2000; 2000JP-0145977.
(UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
WPI: 2002-034733/04.
N-PSDB; ABI99762.
Examining the ischemic condition (e.g. occlusive ischemia) by measuring expression levels of particular genes defined in the specification or by determining the expression profile of a gene group comprising these genes -
Claim 2; Page 2091-2093; 2690pp; English.
The present invention describes a method for examining ischaemic conditions, comprising measuring the expression levels of particular genes (I) in a test sample or determining the expression profile of a gene group in the sample comprising genes selected from (I). The method is useful for examining the ischaemic condition (e.g. compressive ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring expression levels of particular genes (ABI99202 to ABI99912, encoding the protein sequences in ABB57020 to ABB57374) or by determining the expression profile of a gene group comprising these genes. The expression levels or expression profiles produced by these genes are used as an indicator when screening for ischaemic condition-improving drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR primers for a mouse ischaemic condition related sequence, which are used in the exemplification of the present invention.

Alignment Scores:			
Pred. No.:	7.87e-44	Length:	547
Score:	50.00	Matches:	50
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	15.92%	Indels:	0
DB:	23	Gaps:	0
US-09-835-992A-21 (1-994) x ABB57301 (1-547)			
QY	522	ACGCTGTTGATAGAGAGAGATATACATATGAGAGAAAGTGGGTCAATAATCCTAGTGGT	581
Db	329	ThrleuValaspargGlyapsnThrtYrGlyGlylYstPvalIleasnProserGly	348
QY	582	GGACTGATTTCAAGGAGACACCCACTAGGCGCTACAGGTCTTGCTCAGTGTGCAGAACTC	641
Db	349	GlyleuIleSerlysglyHisProleuGlyAlaThrGlyleuAlaGlnCysAlaGluLeu	368
QY	642	TGCTGGCAGCTGAGAGGGGAAACCGGAAAA	671
Db	369	CystPrpGlnleuAArgGlyAlaIaglyLys	378
RESULT 2			
ABB65147			
ID	ABB65147	standard; Protein; 378 AA.	
XX	ABB65147;		
AC			
XX			
DT	26-MAR-2002	(first entry)	
XX			
DE	Drosophila melanogaster	polypeptide SEQ ID NO 22233.	
KW	Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.		
XX			
OS	Drosophila melanogaster.		
XX			
PN	WO200171042-A2.		

XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR N-PSDB; ABL09250.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 22233; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

Alignment Scores:		
Pred. No.:	1.22e-22	Length: 378
Score:	30.00	Matches: 30
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	9.55%	Indels: 0
DB:	22	Gaps: 0
US-09-835-992A-21 (1-994) x ABB65147 (1-378)		
QY	570	AATCCTAGTGTGGAGACTGATTTCAAGGAGACACCCACTAGGCCCTACAGGTCCTGCTCAG 629
Db	308	AsnProSerGlyGlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 327
QY	630	TGTGCAGAACTCTGCTGGCAGCTGAGAGG 659
Db	328	CysAlaGluLeuGlyStrGlnLeuArgGly 337
RESULT 3		
ABB65056		
ID	ABB65056	standard; Protein; 544 AA.
XX		
AC	ABB65056;	
XX		
DT	26-MAR-2002	(first entry)
XX		
DE		
XX		
Drosophila melanogaster polypeptide SEQ ID NO 21960.		
KW		
KW		Drosophila; developmental biology; cell signalling; insecticide;
KW		pharmaceutical.
XX		
OS		
XX		
Drosophila melanogaster.		
XX		
PN		
WO200171042-A2.		
XX		
PD	27-SEP-2001.	
XX		
PF		
23-MAR-2001; 2001WO-US09231.		

XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Ll PWD, Myers EW;
XX WPI; 2001-656860/75.
DR N-PSDB; ABL09159.
DR
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 21.960; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX	Sequence	544	AA;
SQ			
Alignment Scores:			
Pred. No.:		1.18e-22	Length:
Score:		30.00	Matches:
Percent Similarity:		100.00%	Conservative:
Best Local Similarity:		100.00%	Mismatches:
Query Match:		9.55%	Indels:
DB:		22	Gaps:
		0	0

US-09-835-992A-21 (1-994) x ABB65056 (1-544)

QY 570 AATCCTAGTGTGGACTGATTTCAAGGACACCCACTAGGCGCTACAGGTCCTCAG 629
 |||||
 Db 337 AsnProserGlyGlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 356

	RESULT 4
AAE09042	ID AAE09042 standard; Protein; 397 AA.
XX	
AC	AAE09042;
XX	
DT	15-NOV-2001 (first entry)
XX	
DE	Equine Influenza virus H3N8 PelcapA-N-397 protein.
XX	
KW	Equine Influenza virus; el; cold adaptation; temperature sensitivity;
KM	vaccine; pelcapA-N-1217 DNA; PelcapA-N-397 protein.
XX	
OS	Equine Influenza virus H3N8.
XX	
PN	WO200160849-A2.
PD	
PD	23-AUG-2001.
XX	
PF	16-FEB-2001; 2001WO-US05048.
XX	
PR	16-FEB-2000; 2000US-0506286.
XX	
PA	(UYP1-) UNIV PITTSBURGH.

XX Dowling PW, Youngner JS;
PI
XX
XX WPI; 2001-522584/57.
DR N-PSDB; AAD15714.
DR
XX
PT Novel isolated equine influenza virus (wild-type and cold-adapted)
PT proteins and viruses containing nucleic acid molecules encoding the
PT proteins, which are useful for protecting animals from influenza virus
PT infections
XX
PS Claim 5; Page 137-138; 172pp; English.

The patent discloses cold-adapted equine influenza viruses and reassortant influenza A viruses comprising at least one genome segment of such an equine influenza virus, wherein the equine influenza virus genome segment confers at least one identifying phenotype of the cold-adapted equine influenza virus, such as cold adaptation, temperature sensitivity, dominant interference or attenuation. The viruses are useful for protecting animals from diseases caused by influenza viruses. They are also used as vaccines. The present sequence is equine influenza (ei) virus H3N8 Peica (cold adapted) PA-N-397 protein which is encoded by nc1caca-N-1217 DNA.

Seq	Sequence	397	AA;
Alignment Scores:			
Pred. No.:	2.07	Length:	397
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.90%	Indels:	0
DB:	22	Gaps:	0

US-09-835-992A-21 (1-994) x AAE09042 (1-397)

```

OY      337 AACGAGCTGGCAATCAGTCATCAT 311
         |||||
Db      55 ASngluLeuglyvSserValIlele 63

```

PT	RESULT 5
XX	AAE09044
ID	AAE09044 standard; Protein; 716 AA.
XX	
AC	AAE09044;
XX	
DT	15-NOV-2001 (first entry)
XX	
DE	Equine Influenza virus H3N8 Pelcapa-716 protein.
XX	
KW	Equine Influenza virus; e1; cold adaptation; temperature sensitivity;
KW	vaccine; pelcapa-2148 DNA; Pelcapa-716 protein.
XX	
OS	Equine Influenza virus H3N8.
XX	
PN	WO200160849-A2.
XX	
PD	23-AUG-2001.
XX	
PF	16-FEB-2001; 2001WO-US05048.
XX	
PR	16-FEB-2000; 2000US-0506286.
XX	
PA	(UVP1-) UNTV PITTSBURGH.
XX	
PI	Dowling PW, Youngner JS;
XX	
DR	WPI; 2001-522584/57.
DR	N-PSDB; AAD15716.
XX	
PT	Novel isolated equine Influenza virus (wild-type and cold-adapted) proteins and viruses containing nucleic acid molecules encoding the

PT proteins, which are useful for protecting animals from influenza virus
PT infections
XX
PS Claim 5; Page 147-149; 172pp; English.
XX
CC The patent discloses cold-adapted equine influenza viruses and
CC reassortant influenza A viruses comprising at least one genome
CC segment of such an equine influenza virus, wherein the equine
CC influenza virus genome segment confers at least one identifying
CC phenotype of the cold-adapted equine influenza virus, such as
CC cold adaptation, temperature sensitivity, dominant interference
CC or attenuation. The viruses are useful for protecting animals
CC from diseases caused by influenza viruses. They are also used
CC as vaccines. The present sequence is equine influenza (ei) virus
CC H3N8 Peica (cold adapted) PA-716 protein which is encoded by
CC neicaPA-2148 DNA.
XX
SQ Sequence 716 AA;

Alignment Scores:
Pred. No.: 1.95 Length: 716
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.90% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-21 (1-994) x AAE09044 (1-716)

QY 337 AACGAGCTGGCAATCAGTCATCAT 311
Db 55 AsnGlulLeuGllySerValIleIle 63

RESULT 6
AAG20138
ID AAG20138 standard; Protein; 153 AA.
XX
AC AAG20138;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22208.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 26.2 Length: 153
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 21 Gaps: 0

US-09-835-992A-21 (1-994) x AAG20138 (1-153)

QY 597 GGACACCCACTAGCGCTACAGT 620

Db 89 GlyHisProLeuGlyAlaThrGly 96

RESULT 7
AAG39091
ID AAG39091 standard; Protein, 182 AA.
XX
AC AAG39091;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48319.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.:	25 7	Length:	182
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0
DB:	21	Gaps:	0

US-09-835-992A-21 (1-994) x AAG39091 (1-182)

Oy 67 AAATGACACTTGCAAAATTG 90

Db 138 LysLeuAsnThrLeuGlnLysLeu 145

RESULT 8

AAG20137
ID AAG20137 standard; Protein; 187 AA.

XX AAG20137;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 22207.

DE
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 25.6
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.55%
DB: 21
Gaps: 0

US-09-835-992A-21 (1-994) x AAG20137 (1-187)

OY 597 GGACACCCACTAGCGCTACAGGT 620
Db 123 GlyHisProLeucylAlaThrGly 130

RESULT 9
AAG20136
ID AAG20136 standard; protein; 235 AA.
XX
AC AAG20136;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22206.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 25
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.55%
DB: 21
Length: 235
Matches: 8
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-835-992A-21 (1-994) x AAG20136 (1-235)

OY 597 GGACACCCACTAGGCGCTACAGCT 620
Db 171 GLYHISProLeuGLyAlaThrGLy 178

RESULT 10
AAG37797
ID AAG37797 standard; Protein; 241 AA.
XX AAG37797;
AC AAG37797;
XX 18-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46533.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores: 25
Pred. No.: 8.00
Score: 100.00%
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.55%
DB: 21
Gaps: 0

US-09-835-992A-21 (1-994) x AAG37797 (1-241)

OY 597 GGACACCCACTAGGCGCTACAGCT 620
Db 177 GLYHSProlenGlyAlaThrGly 184

RESULT 11
AAG39090
ID AAG39090 standard; Protein; 248 AA.
XX AAG39090;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48318.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.

PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	24-JUN-1999;	99US-0140695.	PR	01-SEP-1999;	99US-0151930.
PR	28-JUN-1999;	99US-0140823.	PR	07-SEP-1999;	99US-0152363.
PR	29-JUN-1999;	99US-0140991.	PR	10-SEP-1999;	99US-0153070.
PR	30-JUN-1999;	99US-0141287.	PR	13-SEP-1999;	99US-0153758.
PR	01-JUL-1999;	99US-0141842.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-0154039.
PR	02-JUL-1999;	99US-0142055.	PR	20-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155486.
PR	09-JUL-1999;	99US-0142920.	PR	24-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	29-SEP-1999;	99US-0156596.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-0157117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-0144085.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144334.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159329.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159331.
PR	21-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159637.
PR	21-JUL-1999;	99US-0144814.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0145086.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160741.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.:	24.9	Length:	248
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0
DB:	21	Gaps:	0

US-09-835-992a-21 (1-994) x AAG39090 (1-248)

OY 67 MAATGACACTTGCMAAATTG 90
|||||
Db 204 LysleuAsnThrLeuGlnLysLeu 211

RESULT 12
AAG37796
ID AAG37796 standard; Protein; 261 AA.

XX AAG37796;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 46532.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.:	24.8	Length:	261
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0
DB:	21	Gaps:	0

US-09-835-992A-21 (1-994) x AAG37796 (1-261)

OY 597 GGACACCCACTAGCGCTACAGT 620
Db 197 GLYHisProLeuGlyAlaThrGly 204

RESULT 13
AAG39054
ID AAG39054 standard; Protein; 284 AA.
XX
AC AAG39054;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48269.
XX
KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX

25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156559.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.:	24.6	Length:	284
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0
DB:	21	Gaps:	0

US-09-835-992A-21 (1-994) x AAG39054 (1-284)

QY 597 GGACACCCACTAGCGCCTACAGGT 620
|||||
Db 225 GlyHisProLeuGlyAlaThrGly 232

RESULT 14
AAG11127
ID AAG11127 standard; Protein; 291 AA.
XX
AC AAG11127;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9727.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores: Pred. No.: 24.5 Length: 291
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 21 Gaps: 0

US-09-835-992a-21 (1-994) x AAG11127 (1-291)

QY 597 GGACACCCACTAGCGCTACAGGT 620
Db 227 GYH1SProlenGlyAlaThrGly 234

RESULT 15
AAG37795
ID AAG37795 standard; Protein; 291 AA.
XX
AC AAG37795;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46531.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.:	24.5	Length:	291
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0
DB:	21	Gaps:	0

US-09-835-992A-21 (1-994) x AAG37795 (1-291)

OY	597	CGACACCCACTAGCGCTACAGCT	620
Db	227	GLYHLPProLeuGlyAlaThrGly	234

Search completed: January 14, 2003, 17:27:24
Job time : 53.9632 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:54 : Search time 13.6606 Seconds
(without alignments)
4281.862 Million cell updates/sec

Title: US-09-835-992A-21
Perfect score: 314
Sequence: 1 ctccaccagtgtgtctctcag.....ccaccctgggtgtgtggat 994

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+n2p.model -DEV-xlp
-O-/cgn2_1/USPTO_spool/US09835992/runat_14012003_161527_4050/app_query.fasta_1.3932
-DB-issued_patents_AA -QFMT-fastlan -SUFFIX-olig.ra1 -MINMATCH-0.1 -LOOPCL-0
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-oligo -TRANS-human40.cdi
-LIST-45 -DOCALIGN-200 -THR_SCORE-quality -THR_MIN-1 -ALIGN-15 -MODE-LOCAL
-OUTFMT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USER-US09835992@cgn_1.1.49_runat_14012003_161527_4050 -NCPU-6 -ICPU-3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -LONGLOG -DEV_TIMEOUT-120
-WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60 -FGAPOP-6 -FGAPEXT-7
-YGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	8	2.5	259	2	US-08-286-819A-40	Sequence 40, Appl
2	8	2.5	259	3	US-08-980-357-40	Sequence 40, Appl
3	8	2.5	384	2	US-08-286-819A-14	Sequence 14, Appl
4	8	2.5	384	3	US-08-980-357-14	Sequence 14, Appl
5	8	2.5	2254	2	US-08-286-819A-28	Sequence 28, Appl
6	8	2.5	2254	3	US-08-980-357-28	Sequence 28, Appl
7	7	2.2	15	3	US-09-045-632-103	Sequence 103, Appl
8	7	2.3	33	5	PCT-US95-12080-4	Sequence 4, Appl
9	7	2.2	119	4	US-08-983-607-34	Sequence 34, Appl
10	7	2.3	188	4	US-08-737-109-14	Sequence 14, Appl
11	7	2.2	219	2	US-08-855-140-4	Sequence 4, Appl
12	7	2.3	260	4	US-08-081-929-10	Sequence 10, Appl

13	7	2.2	268	3	US-08-871-483-11	Sequence 11, Appl
14	7	2.3	270	1	US-08-452-052-1	Sequence 1, Appl
15	7	2.3	274	4	US-08-936-165A-499	Sequence 499, Appl
16	7	2.3	290	4	US-09-230-637-32	Sequence 32, Appl
17	7	2.2	291	4	US-09-134-001C-5669	Sequence 5669, Appl
18	7	2.3	298	3	US-09-025-691-1	Sequence 1, Appl
19	7	2.3	345	1	US-08-134-570-14	Sequence 14, Appl
20	7	2.3	373	2	US-08-015-986A-14	Sequence 14, Appl
21	7	2.3	426	4	US-08-961-083-48	Sequence 48, Appl
22	7	2.2	427	3	US-09-196-857-2	Sequence 2, Appl
23	7	2.2	458	4	US-09-504-358-2	Sequence 2, Appl
24	7	2.2	458	4	US-09-954-314-2	Sequence 2, Appl
25	7	2.2	471	4	US-09-391-104-25	Sequence 25, Appl
26	7	2.3	505	1	US-08-221-750A-5	Sequence 5, Appl
27	7	2.3	540	3	US-08-964-268-6	Sequence 6, Appl
28	7	2.2	585	2	US-08-426-125-2	Sequence 2, Appl
29	7	2.2	585	2	US-08-455-355-2	Sequence 2, Appl
30	7	2.2	610	4	US-09-443-184-55	Sequence 55, Appl
31	7	2.2	785	3	US-08-374-077C-3	Sequence 3, Appl
32	7	2.3	785	4	US-08-895-590-3	Sequence 3, Appl
33	7	2.3	785	4	US-08-539-879A-3	Sequence 3, Appl
34	7	2.3	820	1	US-08-291-896-2	Sequence 2, Appl
35	7	2.2	820	2	US-08-485-278-2	Sequence 2, Appl
36	7	2.2	906	1	US-08-254-573-2	Sequence 2, Appl
37	7	2.2	906	1	US-08-687-379-2	Sequence 2, Appl
38	7	2.2	906	1	US-08-687-379-4	Sequence 4, Appl
39	7	2.2	906	4	US-08-172-332-1	Sequence 1, Appl
40	7	2.2	906	4	US-08-216-326-2	Sequence 2, Appl
41	7	2.2	907	1	US-07-718-575-2	Sequence 2, Appl
42	7	2.2	907	1	US-08-481-206-2	Sequence 2, Appl
43	7	2.2	907	2	US-08-486-269A-2	Sequence 2, Appl
44	7	2.2	1286	6	5206163-1	Patent No. 5206163
45	7	2.2				

ALIGNMENTS

RESULT 1
US-08-286-819A-40
Sequence 40, Application US/08286819A
Patent No. 5871910
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 259 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: E. faecium
STRAIN: BM4147
US-08-286-819A-40

Alignment Scores:
Pred. No.: 7.98 Length: 259
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x US-08-286-819A-40 (1-259)

QY 58 ATGAACAAAATTGAACACTTTG 81
|||||
Db 5 MetGlutInIlyslLeuAsnThrIeu 12

RESULT 2
US-08-980-357-40
Sequence 40, Application US/08980357
Patent No. 6013508
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 259 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: E. faecium
STRAIN: BM4147
US-08-980-357-40

Alignment Scores:
Pred. No.: 7.98 Length: 259
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 3 Gaps: 0

US-09-835-992A-21 (1-994) x US-08-980-357-40 (1-259)

QY 58 ATGAACAAAATTGAACACTTTG 81
|||||
Db 5 MetGlutInIlyslLeuAsnThrIeu 12

RESULT 3
US-08-286-819A-14
Sequence 14, Application US/08286819A
Patent No. 5871910
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-286-819A-14

Alignment Scores:
Pred. No.: 7.63 Length: 384
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x US-08-286-819A-14 (1-384)

QY 58 ATGGAACAAAATTGAACACTTTG 81
|||||
Db 130 MetGlulnLysLeuAsnThrLeu 137

RESULT 4
US-08-980-357-14
Sequence 14, Application US/08980357
Patent No. 6013508
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPROTEINS, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia

COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-980-357-14

Alignment Scores:
Pred. No.: 7.63 Length: 384
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 3 Gaps: 0

US-09-835-992A-21 (1-994) x US-08-980-357-14 (1-384)

QY 58 ATGGAACAAAATTGAACACTTTG 81
|||||
Db 130 MetGlulnLysLeuAsnThrLeu 137

RESULT 5
US-08-286-819A-28
Sequence 28, Application US/08286819A
Patent No. 5871910
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPROTEINS, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400

CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 2254 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-286-819A-28

Alignment Scores:
Pred. No.: 6.27 Length: 2254
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x US-08-286-819A-28 (1-2254)

OY 58 ATGGAACAAATGACACTTTG 81
|||||
Db 808 MetGlulnLysLeuAsnThrLeu 815

RESULT 6
US-08-980-357-28
Sequence 28, Application US/08980357
Patent No. 6013508
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPROTEINS, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 2254 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-980-357-28

Alignment Scores:
Pred. No.: 6.27 Length: 2254
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 3 Gaps: 0

US-09-835-992A-21 (1-994) x US-08-980-357-28 (1-2254)

OY 58 ATGGAACAAATGACACTTTG 81
|||||
Db 808 MetGlulnLysLeuAsnThrLeu 815

RESULT 7
US-09-045-632-103
Sequence 103, Application US/09045632
Patent No. 6001575
GENERAL INFORMATION:
APPLICANT: Huganir, Richard L.
APPLICANT: Dong, Hualing
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
TITLE OF INVENTION: GRIP-RELATED MOLECULES
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-045-632-103

Alignment Scores:
Pred. No.: 115 Length: 15
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 3 Gaps: 0

US-09-835-992A-21 (1-994) x US-09-045-632-103 (1-15)

OY 603 CCACTAGGCGCTACAGTCTT 623
Db 9 ProleuglyAlathrglyLeu 15

RESULT 8
PCT-US95-12080-4
Sequence 4, Application PC/TUS9512080
GENERAL INFORMATION:
APPLICANT: Children's Medical Center Corporaton
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12080
FILING DATE:
CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-815-8795

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-12080-4

Alignment Scores:
Pred. No.: 105 Length: 33
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 5 Gaps: 0

US-09-835-992A-21 (1-994) x PCT-US95-12080-4 (1-33)

OY 194 CAAAAATCAAAACCTTCCTT 174
Db 19 GlnlysllelyAsnphpe 25

RESULT 9
US-08-983-607-34
Sequence 34, Application US/08983607
Patent No. 6140470
GENERAL INFORMATION:
APPLICANT: Alan Garen
TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-
TITLE OF INVENTION: bodies
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Department of Molecular Biophysics
ADDRESSEE: and Biochemistry, Yale University
STREET: 266 Whitney Avenue
CITY: New Haven
STATE: Connecticut
COUNTRY: United States of America
ZIP: 06520-8114
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" 1.44 MB diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processing
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/983,607
FILING DATE: April 27, 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB96/01032
FILING DATE: June 28, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mary M. Krinsky
REGISTRATION NUMBER: 32423
REFERENCE/DOCKET NUMBER: OCR-679
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-773-9544
TELEFAX: 203-773-1183
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: polypeptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens (melanoma patient immu-
ORGANISM: nized with autologous tumor cells)
INDIVIDUAL ISOLATE: peripheral blood lymphocytes
IMMEDIATE SOURCE:
LIBRARY: DM414 scfv antibodies obtained from

LIBRARY: fuses fusion phage construct
CLONE: V373
FEATURE:
NAME/KEY: heavy chain
OTHER INFORMATION: Xaa at position 3 is Gln or
OTHER INFORMATION: Xaa and Xaa at position 19 is Xaa or Ser
US-08-983-607-34

Alignment Scores:
Pred. No.: 91 Length: 119
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-21 (1-994) x US-08-983-607-34 (1-119)

OY 564 GTCATAAATCCTAGTGGA 584
|||||
Db 48 ValIleAsnProSerGlyGly 54

RESULT 10
US-08-737-109-14
; Sequence 14, Application US/08737109
; Patent No. 6455688
; GENERAL INFORMATION:
; APPLICANT: SLABAS, Antoni Ryszard
; APPLICANT: ELBOROUGH, Kieran Michael
; APPLICANT: BRIGHT, Simon William Jonathan
; APPLICANT: FENTEM, Philip Anthony
; TITLE OF INVENTION: Plant Gene Specifying Acetyl Coenzyme A
; TITLE OF INVENTION: Carboxylase and Transformed Plants Containing Same
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,109
; FILING DATE: 21-OCT-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/00846
; FILING DATE: 02-MAY-1994
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; IMMEDIATE SOURCE:
; CLONE: pK1U81
US-08-737-109-14

Alignment Scores:
Pred. No.: 86.5 Length: 188
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-21 (1-994) x US-08-737-109-14 (1-188)

OY 269 CTTCTGTACAAATGCTTCACT 249
|||||
Db 154 LeuLeuTyrIlyscyspHeThr 160

RESULT 11
US-08-855-140-4
; Sequence 4, Application US/08855140
; Patent No. 5854022
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: NEW TRANSMEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/855,140
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0296 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 219 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1279546
US-08-855-140-4

Alignment Scores:
Pred. No.: 85 Length: 219
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x US-08-855-140-4 (1-219)

OY 61 GAACAATAATGACACTTG 81
|||||
Db 107 GluGlnIlysLeuAsnThrIleu 113

RESULT 12
US-08-081-929-10
; Sequence 10, Application US/08081929

Patent No. 6160090
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Barnea, Gilad
APPLICANT: Grumet, Martin H.
APPLICANT: Margolis, Richard U.
TITLE OF INVENTION: A NEW CLASS OF RPTASES: THEIR
STRUCTURAL DOMAINS AND LIGANDS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/081,929
FILING DATE: 23-JUN-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30742
REFERENCE/DOCKET NUMBER: 7683-041-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 260 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-081-929-10

Alignment Scores:
Pred. No.: 83.4
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.26%
DB: 4

US-09-835-992a-21 (1-994) x US-08-081-929-10 (1-260)

OY 942 CTTGAGCGCAAAATTCN 922
|||||
Db 114 LeuGluglyGlnLysphePro 120

RESULT 13
US-08-871-483-11
Sequence 11, Application US/08871483
Patent No. 6083715
GENERAL INFORMATION:
APPLICANT: Georgiou, George
APPLICANT: Oul, Ji
APPLICANT: Bessette, Paul
APPLICANT: Swartz, James
TITLE OF INVENTION: METHODS FOR PRODUCING HETEROLOGOUS
TITLE OF INVENTION: DISULFIDE BOND-CONTAINING PLOYPEPTIDES IN BACTERIAL CELLS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas

COUNTRY: United States
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,483
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UTSB:614
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-871-483-11

Alignment Scores:
Pred. No.: 83.1
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.23%
DB: 3

US-09-835-992a-21 (1-994) x US-08-871-483-11 (1-268)

OY 231 GCAGCAGCATTTGGCAGT 251
|||||
Db 183 AlaAlaAlaIleuAlaSer 189

RESULT 14
US-08-452-052-1
Sequence 1, Application US/08452052
Patent No. 5766922
GENERAL INFORMATION:
APPLICANT: Peles, Elior
TITLE OF INVENTION: FUNCTIONAL LIGANDS FOR THE AXONAL CELL
TITLE OF INVENTION: RECOGNITION MOLECULE CONTACTIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,052
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-075
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:29:25 ; Search time 9.51637 Seconds
(without alignments)
4052.943 Million cell updates/sec

Title: US-09-835-992A-21
Perfect score: 314
Sequence: 1 ctccaccagtcgtctctcag.....ccaccctgggtgtgtgat 994

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 118974 seqs, 19401057 residues

Word size: 1

Total number of hits satisfying chosen parameters: 214280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09835992/runat_14012003_161528_4118/app_query.fasta_1.3932
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=olig.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09835992_@CGN_1_1_28_@runat_14012003_161528_4118
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	8	2.5	143	12 US-10-062-254-368 Sequence 368, App
2	8	2.5	359	9 US-10-012-055-4 Sequence 4, Appl
3	8	2.5	397	10 US-09-815-242-12461 Sequence 12461, A
4	7	2.2	42	10 US-09-864-761-41400 Sequence 41400, A

C	5	7	2.2	52	10	US-09-864-761-35057	Sequence 35057, A
C	6	7	2.3	54	10	US-09-864-761-46413	Sequence 46413, A
C	7	7	2.3	60	10	US-09-864-761-45360	Sequence 45360, A
C	8	7	2.3	64	9	US-09-796-692-1638	Sequence 1638, Ap
C	9	7	2.3	64	9	US-09-796-692-1978	Sequence 1978, Ap
C	10	7	2.2	71	12	US-10-078-929-174	Sequence 174, App
C	11	7	2.2	74	10	US-09-864-761-42412	Sequence 42412, A
C	12	7	2.2	74	12	US-10-078-929-176	Sequence 176, App
C	13	7	2.3	87	9	US-09-796-692-1515	Sequence 1515, Ap
C	14	7	2.3	87	9	US-09-796-692-1610	Sequence 1610, Ap
C	15	7	2.3	87	9	US-09-796-692-1745	Sequence 1745, Ap
C	16	7	2.3	87	9	US-09-796-692-1758	Sequence 1758, Ap
C	17	7	2.3	87	9	US-09-796-692-1942	Sequence 1942, Ap
C	18	7	2.3	87	9	US-09-796-692-1964	Sequence 1964, Ap
C	19	7	2.3	87	9	US-09-796-692-2027	Sequence 2027, Ap
C	20	7	2.3	87	9	US-09-796-692-2040	Sequence 2040, Ap
C	21	7	2.2	132	10	US-09-811-737-10	Sequence 10, Appl
C	22	7	2.3	149	10	US-09-864-761-43779	Sequence 43779, A
C	23	7	2.2	159	9	US-09-764-868-960	Sequence 960, App
C	24	7	2.2	159	10	US-09-780-017-2	Sequence 2, Appl1
C	25	7	2.2	173	10	US-09-925-297-688	Sequence 688, App
C	26	7	2.2	239	9	US-09-738-626-3774	Sequence 3774, Ap
C	27	7	2.2	260	10	US-09-811-737-16	Sequence 16, Appl
C	28	7	2.3	260	12	US-10-000-954-10	Sequence 10, Appl
C	29	7	2.3	274	10	US-09-939-980-499	Sequence 499, App
C	30	7	2.3	294	10	US-09-815-242-5886	Sequence 5886, Ap
C	31	7	2.3	298	10	US-09-748-033-1	Sequence 1, Appl1
C	32	7	2.3	299	10	US-09-815-242-12904	Sequence 12904, A
C	33	7	2.3	299	10	US-09-815-242-13086	Sequence 13086, A
C	34	7	2.2	358	10	US-09-815-242-5248	Sequence 5248, Ap
C	35	7	2.2	358	10	US-09-815-242-12363	Sequence 12363, A
C	36	7	2.2	426	10	US-09-765-272-48	Sequence 48, Appl
C	37	7	2.2	458	10	US-09-954-314-2	Sequence 2, Appl1
C	38	7	2.3	471	10	US-09-801-196-21	Sequence 21, Appl
C	39	7	2.3	471	10	US-09-801-196-32	Sequence 32, Appl
C	40	7	2.2	482	12	US-10-078-929-180	Sequence 180, App
C	41	7	2.3	545	10	US-09-908-988B-4	Sequence 4, Appl1
C	42	7	2.3	548	10	US-09-515-806-3	Sequence 3, Appl1
C	43	7	2.3	604	10	US-09-925-301-1026	Sequence 1026, Ap
C	44	7	2.2	610	9	US-09-992-598-113	Sequence 113, App
C	45	7	2.2	610	9	US-09-989-293A-113	Sequence 113, App

ALIGNMENTS

RESULT 1
US-10-062-254-368
; Sequence 368, Application US/10062254
; Patent No. US20020138882A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fang, Yiwen
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jian-Ming
; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping
; APPLICANT: Odell, Joan
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; APPLICANT: Zheng, Peizhong
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/062, 254
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/630, 346
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30

; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/157287
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/169767
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/171054
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/172958
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/171515
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/173535
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 368
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-062-254-368

Alignment Scores:
Pred. No.: 4.25 Length: 143
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 12 Gaps: 0

US-09-835-992A-21 (1-994) x US-10-062-254-368 (1-143)

QY 760 TTCCCGAAGCCGCGAGTTCCTTT 783
|||||
Db 3 PheProgluAlaAlaSerSerPhe 10

RESULT 2

US-10-012-055-4
; Sequence 4, Application US/10012055
; Patent No. US20020164750A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: 56634, A NOVEL HUMAN PHOSPHATIDYLINOSITOL 4-PHOSPHATE 5-KINASE
; FILE REFERENCE: 10448-114001
; CURRENT APPLICATION NUMBER: US/10/012,055
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/248,325
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
US-10-012-055-4

Alignment Scores:
Pred. No.: 3.92 Length: 359
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 9 Gaps: 0

US-09-835-992A-21 (1-994) x US-10-012-055-4 (1-359)

QY 337 TTGAAGAAAAAGCATTTATTAATA 360

|||||
Db 152 LeuYsLysLysAlaLeuLeuLys 159

RESULT 3

US-09-815-242-12461
; Sequence 12461, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12461
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12461

Alignment Scores:
Pred. No.: 3.89 Length: 397
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-21 (1-994) x US-09-815-242-12461 (1-397)

QY 597 GGACACCCACTAGCGCTACAGT 620
|||||
Db 350 GlyHisProLeuGlyAlaThrGly 357

RESULT 4

US-09-864-761-41400
; Sequence 41400, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04

```

; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 4917
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41400

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC027273.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
US-09-864-761-41400

```

Alignment Scores:		
Pred. No.:	53	42
Score:	7.00	7
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	2.23%	Indels: 0
DB:	10	Gaps: 0

US-09-835-992A-21 (1-994) x US-09-864-761-41400 (1-42)

OY	337	TTCAGCAAAAGCATTATTA	357
Dd	11	leuLysLSlySLalaleuLcu	17

RESULT 5
US-09-864-761-35057
; Sequence 35057, Application US/09864761
; Patent NO. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.

```

; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761

```

```

CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 35057

```

```

1  LENGTH: 52
2  TYPE: PRT
3  ORGANISM: Homo sapiens
4  FEATURE:
5  OTHER INFORMATION: MAP TO AL080245.13
6  OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 1.3
7  OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 1.4
8  OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL - 1.7
9  OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 1.4
10 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 1.7
11 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL - 2.1
12 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 1.3
13 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 1.9
14 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 1.9
15 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 1
16 OTHER INFORMATION: SWISSPROT HIT: P54697, EVALUE 1.20e-02
17 OTHER INFORMATION: EST_HUMAN HIT: AA993492.1, EVALUE 3.00e-25
18
19 US-09-864-761-35057

```

Alignment Scores:	
Pred. No.:	52.1
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	2.23%
DB:	10
Length:	52
Matches:	7
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-835-992A-21 (1-994) x US-09-864-761-35057 (1-52)

QY 455 TTGCTTTCTACCAAGAACT 475

DB 36 LeuLeuphetyrGlnArgThr 42

RESULT 6

US-09-864-761-46413

; Sequence 46413, Application US/09864761
; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-x-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 46413

; LENGTH: 54

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC020897.3

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3

; OTHER INFORMATION: SWISSPROT HIT: O32869, EVALUE 7.00e-01

; OTHER INFORMATION: EST_HUMAN HIT: BF436304.1, EVALUE 3.70e+00

US-09-864-761-46413

Alignment Scores:

Pred. No.: 51.9

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.26%

DB: 10

US-09-835-992A-21 (1-994) x US-09-864-761-46413 (1-54)

QY 782 AAGAACTGGCGCTTCGGG 762

DB 12 LysgluLeuAlaIasercgly 18

RESULT 7

US-09-864-761-45360

; Sequence 45360, Application US/09864761
; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecmica-x-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 45360

; LENGTH: 60

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC013626.3

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EST HUMAN HIT: F06450.1, EVALUE 4.00e-26
OTHER INFORMATION: SWISSPROT HIT: P36607, EVALUE 1.00e-08
US-09-864-761-45360

Alignment Scores:
Pred. No.: 51.4 Length: 60
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-21 (1-994) x US-09-864-761-45360 (1-60)

OY 657 CTCTCAGCTGCAGCAGCTT 637
|||||
Db 49 LeuserAlaAlaSerArgVal 55

RESULT 8

US-09-796-692-1638
Sequence 1638, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1638
LENGTH: 64
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(64)
OTHER INFORMATION: Xaa = Any amino acid
US-09-796-692-1638

Alignment Scores:
Pred. No.: 51.1 Length: 64
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.26% Indels: 0
DB: 9 Gaps: 0

US-09-835-992A-21 (1-994) x US-09-796-692-1638 (1-64)

OY 663 CTCCCTCTCAGCTGCAGC 643
|||||
Db 24 LeuproLeuserAlaAlaSer 30

RESULT 9

US-09-796-692-1978
Sequence 1978, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1978
LENGTH: 64
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(64)
OTHER INFORMATION: Xaa = Any amino acid
US-09-796-692-1978

Alignment Scores:
Pred. No.: 51.1 Length: 64
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 9 Gaps: 0

US-09-835-992A-21 (1-994) x US-09-796-692-1978 (1-64)

OY 663 CTCCCTCTCAGCTGCAGC 643
|||||
Db 24 LeuproLeuserAlaAlaSer 30

RESULT 10

US-10-078-929-174
; Sequence 174, Application US/10078929
; Patent No. US20020152497A1
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Antoni
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Sakai, Hajime
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Odell, Joan T.
; APPLICANT: Meyers, Blake
; APPLICANT: Thorpe, Catherine
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in
; TITLE OF INVENTION: Stress Response
; FILE REFERENCE: B01357 US NA
; CURRENT APPLICATION NUMBER: US/10/078,929
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/566,394
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 60/133038
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133042
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133427
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133437
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133428
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133438
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133436
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/137667
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 174
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Glycine max
US-10-078-929-174

Alignment Scores:
Pred. No.: 50.7
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.23%
DB: 12
Gaps: 0

US-09-835-992A-21 (1-994) x US-10-078-929-174 (1-71)
QY 780 CTTTAGAAGCTCATCAAAATT 800
Db 43 LeuLeuGluLeuLeuLeuLeuLysile 49

RESULT 11
US-09-864-761-42412
; Sequence 42412, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42412
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO 297197.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EST_HUMAN HIT: AA099217.1, EVALUATE 7.00e-19
; OTHER INFORMATION: SWISSPROT HIT: Q92793, EVALUATE 6.40e+00
US-09-864-761-42412

Alignment Scores:
Pred. No.: 50.5
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.23%
DB: 10
Gaps: 0

US-09-835-992A-21 (1-994) x US-09-864-761-42412 (1-74)
QY 809 TTCCAACCAAGCTGTGCAAGT 829
Db 67 PheGlnProSerSerAlaSer 73

RESULT 12
US-10-078-929-176
; Sequence 176, Application US/10078929
; Patent No. US20020152497A1


```
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Antoni
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Falco, Savergio Carl
; APPLICANT: Sakai, Hajime
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Odell, Joan T.
; APPLICANT: Meyers, Blake
; APPLICANT: Thorpe, Catherine
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in
; TITLE OF INVENTION: Stress Response
; FILE REFERENCE: B1357 US NA
; CURRENT APPLICATION NUMBER: US/10/078,929
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/566,394
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 60/133038
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133042
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133427
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133437
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133428
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133438
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133436
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/137667
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 176
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-078-929-176

Alignment Scores:
Pred. No.: 50.5      Length: 74
Score: 7.00         Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 2.23%              Indels: 0
DB: 12                      Gaps: 0

US-09-835-992A-21 (1-994) x US-10-078-929-176 (1-74)
QY 780 CTTTGAAGCTCATCAAAAT 800
Db 45 LeuLeuGlulLeuIleLysIle 51

RESULT 13
US-09-796-692-1515
; Sequence 1515, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
```

```
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1515
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-796-692-1515

Alignment Scores:
Pred. No.: 49.8      Length: 87
Score: 7.00         Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 2.26%              Indels: 0
DB: 9                      Gaps: 0

US-09-835-992A-21 (1-994) x US-09-796-692-1515 (1-87)
QY 663 CTTCCCTCTCAGTCGACG 643
Db 47 LeuProLeuSerAlaAlaSer 53

RESULT 14
US-09-796-692-1610
; Sequence 1610, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-07-14
```

;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1610
;; LENGTH: 87
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-796-692-1610

Alignment Scores:
Pred. No.: 49.8 Length: 87
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 9 Gaps: 0

US-09-835-992A-21 (1-994) x US-09-796-692-1610 (1-87)

OY 663 CTTCCCTCTCAGCTGCCAGC 643
Db 47 LeuProleuserAlaAlaser 53

RESULT 15
US-09-796-692-1745
; Sequence 1745, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1745
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1745

Alignment Scores:

Pred. No.: 49.8 Length: 87
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 9 Gaps: 0
US-09-835-992A-21 (1-994) x US-09-796-692-1745 (1-87)
OY 663 CTTCCCTCTCAGCTGCCAGC 643
Db 47 LeuProleuserAlaAlaser 53

Search completed: January 14, 2003, 18:08:20
Job time : 11.5164 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:45 ; Search time 33.3073 Seconds
(without alignments)
5737.940 Million cell updates/sec

Title: US-09-835-992A-21
Perfect score: 314
Sequence: 1 ctcaaccagtgctcctcag.....ccaccctggggtgtgggat 994

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565536

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:
-MODEL-frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09835992/runat_14012003_161526_4038/app-query.fasta_1.3932
-DB-PIR_73 -QFMT=fastan -SUFFIX=olig.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09835992_@CGN_1_1_192_@runat_14012003_161526_4038 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	223	71.0	547	2	I38205		sterol carrier pro
2	84	26.8	546	2	B40407		sterol carrier pro
3	50	15.9	547	2	JU0157		sterol carrier pro
4	49	15.6	547	2	A39368		sterol carrier pro
5	33	10.5	547	2	S34744		sterol carrier pro
6	24	7.6	401	2	E70884		probable lipid-tra
7	22	7.0	412	2	T27202		hypothetical prote
8	10	3.2	387	2	B90436		hypothetical prote
9	9	2.9	246	2	T45470		ATP-binding ABC tr
10	9	2.9	351	2	A96559		hypothetical prote
11	9	2.9	383	2	G69205		lipid-transfer pro
12	8	2.5	143	2	A40015		sterol carrier pro
13	8	2.5	231	2	T19267		hypothetical prote
14	8	2.5	267	2	T35889		hypothetical prote

15	8	2.5	308	2	S58137	gene 7 protein - p
c	16	2.6	314	2	T05993	probable peroxidase
c	17	2.6	375	1	S66272	alcohol dehydrogen
18	8	2.5	379	2	C83278	probable acyl-CoA
19	8	2.5	383	2	E84249	3-ketoacyl-CoA thl
20	8	2.5	384	2	B41838	two-component sens
21	8	2.5	386	2	B70604	probable FadA6 pro
22	8	2.5	387	2	F69304	3-ketoacyl-CoA thl
23	8	2.5	388	2	H69370	3-ketoacyl-CoA thl
24	8	2.5	391	2	B69252	3-ketoacyl-CoA thl
25	8	2.5	394	2	B89786	hypothetical prote
26	8	2.5	395	2	H90421	hypothetical prote
27	8	2.5	398	1	B64383	molybdenum cofacto
28	8	2.5	400	2	A87427	thiolase family pr
29	8	2.5	401	2	E83427	probable acyl-CoA
30	8	2.5	401	2	D87258	fatty oxidation co
31	8	2.5	402	2	B97420	probable acyl-CoA
32	8	2.5	402	2	AC2638	acyl-CoA thiolase
33	8	2.5	403	2	C70815	probable beta-keto
34	8	2.5	403	2	E87179	probable beta-keto
35	8	2.5	404	2	T35428	probable acetyl co
36	8	2.5	404	2	T35256	probable thiolase
37	8	2.5	407	2	E87634	thiolase family pr
38	8	2.5	412	2	B70582	probable lipid-tra
40	8	2.5	414	2	T52165	acetyl-CoA C-acylt
41	8	2.5	417	2	AG3571	acetyl-CoA C-acylt
42	8	2.5	426	2	T01790	protoporphyrin IX
43	8	2.5	430	2	S57792	acetyl-CoA C-acylt
44	8	2.5	436	2	D84782	probable proline t
c	45	2.6	440	2	T24837	hypothetical prote

ALIGNMENTS

RESULT 1
I38205
sterol carrier protein-X/sterol carrier protein-2 - human
C:Species: Homo sapiens (man)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 08-Sep-2000
C:Accession: I38205
R:Ohba, T.; Rennett, H.; Pfeiffer, S.M.; He, Z.; Yamamoto, R.; Holt, J.A.; Billheimer, Genomics 24, 370-374, 1994
A:Title: The structure of the human sterol carrier protein X/sterol carrier protein 2
A:Reference number: I38205; MUID:95213031; PMID:7698762
A:Accession: I38205
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-547 <RES>
A:Cross-references: EMBL:U11313; NID:9532077; PID:9532079
A:Note: nucleotide sequence is not complete
C:Genetics:
A:Gene: GDB:SCP2
A:Cross-references: GDB:126877; OMIM:184755
A:Map position: 1p32-1p32
A:Introns: 23/3; 43/1; 67/1; 111/1; 132/3; 175/1; 196/2; 225/2; 275/3; 325/1; 361/1;

Alignment Scores:
Pred. No.: 5.78e-224 Length: 547
Score: 223.00 Matches: 223
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 71.02% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x I38205 (1-547)
Oy 3 CACCCAGTGTCTCTCAGATGTTGGGTATGCTGAAAAGACATATGGAATAATGGA 62
Db 156 HSPVVALAIAProGlnMetPheGlyTYrAlaGlyLysGluHisMetCylLysTYrGly 175
Oy 63 ACMAAAATGAACACTTTGCCAAAATATGATGGAATAATCATAAACATTCAGTTAATAC 122

Db 176 ThrLysIleGluHisPheAlaLysIleGlyTrpLysAsnHisLysHisSerValAsnAsn 195

QY 123 CCGTATCCAGTTCACAGATGAATACAGTTAGATGAAGTGGCATCTAAAGAAGTT 182

Db 196 ProTyrSerGlnPheGlnAspGluTyrSerLeuAspGluValMetAlaSerLysGluVal 215

QY 183 TTTGATTTTGTGACTATCTTACAATGTGTCCCACTTCAGATGGTGGTCCAGCAAT 242

Db 216 PheAspPheLeuThrIleLeuGlnCysCysProThrSerAspGlyAlaAlaAlaIle 235

QY 243 TTGGCCAGTGAAGCATTTGTACAGAGATGGCCCTGCAATCCAAAGCTGTGGAATTTTG 302

Db 236 LeuAlaSerGluAlaPheValGlnLysTyrGlyLeuGlnSerLysAlaValGlnIleLeu 255

QY 303 GCACAGAAATGATGACTGATTTGCCAAGCTCGTTGAGAGAAAAAGCATTTAAATG 362

Db 256 AlaGlnGluMetThrAspLeuProSerSerPheGlnGluLysSerIleIleLysMet 275

QY 363 GTTGGCTTTGATATGAGTAAAGAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGACA 422

Db 276 ValGlyPheAspMetSerLysGlnAlaAlaArgLysCysTyrGlnLysSerGlyLeuThr 295

QY 423 CCAATGATATTGACGTAATAGAACTTCACGATTTGCTTTTACCAAGAACTCCTACT 482

Db 296 ProAsnAspIleAspValIleGlnLeuHisAspCysPheSerThrAsnGluLeuLeuThr 315

QY 483 TATGAAGCACTGGGACTCTGTCCAGAGACAAAGGTCACACGCTGGTGTATAGAGAGAT 542

Db 316 TyrGluAlaLeuGlyLeuCysProGlnGlyGlnGlyAlaThrLeuValAspArgGlyAsp 335

QY 543 AATACATATGGAGGAAAGTGGGTCAATCCCTAGTGGTGGACTGATTTCAAGGACAC 602

Db 336 AsnThrTyrGlyGlyLysTrpValIleAsnProSerGlyLeuIleSerLysGlyHis 355

QY 603 CCACTAGGCGCTACAGGTCTGTCTCAGTGTGCAGAACTCTGTGGCAGCTGAGAGGGAA 662

Db 356 ProLeuGlyAlaThrGlyLeuAlaGlnCysAlaGluLeuCysTrpGlnLeuArgGlyGlu 375

QY 663 GCCGGA AAA 671

Db 376 AlaGlyLys 378

RESULT 2

B40407

sterol carrier protein 2-related form, 58.85K - human

N:Alternate names: nonspecific lipid transfer protein

N:Contains: sterol carrier protein 2 precursor; sterol carrier protein 2-related form, 3

C:Species: Homo sapiens (man)

C:Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 05-Nov-1999

C:Accession: B40407; A40407; A39010; I54246

R:He, Z.; Yamamoto, R.; Furth, E.E.; Schantz, L.J.; Naylor, S.L.; George, H.; Billheimer

DNA Cell Biol. 10, 559-569, 1991

A:Title: cDNAs encoding members of a family of proteins related to human sterol carrier

A:Reference number: A40407; MUID:92029618; PMID:1718316

A:Accession: B40407

A:Molecule type: mRNA

A:Residues: 1-546 <HEA>

A:Cross-references: GB:M75884

A:Note: this form appears to be present in the matrix of peroxisomes

A:Accession: A40407

A:Molecule type: mRNA

A:Residues: 258-546 <HE2>

A:Cross-references: GB:M75883

R:Yamamoto, R.; Kallen, C.B.; Babalola, G.O.; Rennert, H.; Billheimer, J.T.; Strauss III

Proc. Natl. Acad. Sci. U.S.A. 88, 463-467, 1991

A:Title: Cloning and expression of a cDNA encoding human sterol carrier protein 2.

A:Reference number: A39010; MUID:91110550; PMID:1703300

A:Accession: A39010

A:Molecule type: mRNA

A:Residues: 404-546 <YAM>

A:Cross-references: GB:M55421

A:Note: this mRNA contains an in-frame stop codon in the 5'-untranslated region

A:Note: this form appears to be present in cytosol, mitochondria, microsomes, and on the

R:Yamamoto, R.

Hokkaido Igaku Zasshi 67, 839-848, 1992

A:Title: [localization of human sterol carrier protein 2 gene and cDNA expression in

A:Reference number: I54246; MUID:93131254; PMID:1483685

A:Accession: I54246

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 404-546 <RES>

A:Cross-references: GB:S52450; NID:g263550; PIDN:AAB24921.1; PID:g263551

C:Genetics:

A:Gene: GDB:SCP2

A:Cross-references: GDB:126877; OMIM:184755

A:Map position: 1p32-1p32

C:Keywords: alternative initiators; alternative splicing; peroxisome

F:404-546/Product: sterol carrier protein 2 precursor #status predicted <PRE>

F:404-423/Domain: signal sequence #link PRE #status predicted <SIG>

F:424-546/Product: sterol carrier protein 2 #status predicted <MAT>

F:544-546/Region: peroxisome/glyoxysome location signal #status atypical

Alignment Scores:

Pred. No.: 8.34e-79 Length: 546

Score: 84.00 Matches: 84

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 26.75% Indels: 0

DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x B40407 (1-546)

QY 3 CACCAGTTGCTCCTCAGATGTTGGGTATGCTGGAAAAAGACATATGAAAAATATGCA 62

Db 155 HisProValAlaProGlnMetPheGlyTyrAlaGlyLysGlnHisMetGluLysTyrGly 174

QY 63 ACAAAATTTGAACACTTTGCCAAAAATTTGGATGAAAAATCATAAACATTGAGTTAATAAC 122

Db 175 ThrLysIleGluHisPheAlaLysIleGlyTrpLysAsnHisLysHisSerValAsnAsn 194

QY 123 CCGTATCCAGTTCACAGATGAATACAGTTTAGATGAAGTGAATGCATCTAAAGAAGTT 182

Db 195 ProTyrSerGlnPheGlnAspGluTyrSerLeuAspGluValMetAlaSerLysGluVal 214

QY 183 TTTGATTTTGTGACTATCTTACAATGTGTCCCACTTCAGATGGTGGTCCAGCAAT 242

Db 215 PheAspPheLeuThrIleLeuGlnCysCysProThrSerAspGlyAlaAlaAlaIle 234

QY 243 TTGGCCAGTGAA 254

Db 235 LeuAlaSerGlu 238

RESULT 3

JU0157

sterol carrier protein x - mouse

N:Contains: sterol carrier protein 2

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999

C:Accession: JU0157

R:Seedorf, U.; Raabe, M.; Assmann, G.

Gene 123, 165-172, 1993

A:Title: Cloning, expression and sequences of mouse sterol-carrier protein-x-encoding

A:Reference number: JU0157; MUID:93154580; PMID:8428655

A:Accession: JU0157

A:Molecule type: mRNA

A:Residues: 1-547 <SEE>

A:Cross-references: GB:M91458; NID:g293793; PIDN:AAA40098.1; PID:g293794

C:Genetics:

A:Gene: SCP-x

C:Keywords: carrier protein; mitochondrion

F:1-547/Product: sterol carrier protein x #status predicted <STX>

F:404-547/Product: sterol carrier protein 2 #status predicted <ST2>

Alignment Scores:

Pred. No.: 2.68e-43 Length: 547

Score: 50.00 Matches: 50

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.92% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-21 (1-994) x JU0157 (1-547)

OY 522 ACGCTGCTTGATAGACAGATATACATATGAGAGAAAGTGGCTCATTAATCCTAGTGGT 581
|||||
Db 329 ThrLeuValAspArgGlyAspAsnThrTyrglyGlyserPValIleAsnProsergly 348
OY 582 GCACTGATTTCAAGGACACCCACTAGGCGCTACAGGTCCTCAGTGTGCAGACTC 641
|||||
Db 349 GlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGlnCysAlaGluLeu 368
OY 642 TGGTGGCAGCTGAGAGGGGAAGCCGGA AAA 671
|||||
Db 369 CysTrpGlnLeuArgGlyGlyAlaGlyLys 378

RESULT 4

A39368 sterol carrier protein 2 precursor - rat

C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 01-Dec-2000
C:Accession: A39368; A39054; A29366; A27661; S17842; A34584; A34635
R:Morl, T.; Tsukamoto, T.; Morl, H.; Tashiro, Y.; Fujiki, Y.
Proc. Natl. Acad. Sci. U.S.A. 88, 4338-4342, 1991
A:Title: Molecular cloning and deduced amino acid sequence of nonspecific lipid transfer
sequence of nonspecific lipid transfer protein as its C-terminal part.

A:Reference number: A39368; MUID:91239563; PMID:2034675

A:Accession: A39368

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-547 <MOR>

A:Cross-references: GB:M62763; NID:g202552; PIDN:AAA40622.1; PID:g202553

R:Seedorf, U.; Assmann, G.

J. Biol. Chem. 266, 630-636, 1991

A:Title: Cloning, expression, and nucleotide sequence of rat liver sterol carrier protein

A:Reference number: A39054; MUID:91093192; PMID:1985920

A:Accession: A39054

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 405-547 <SEE>

A:Cross-references: GB:J05716

A:Accession: B39054

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-11, 'P', 13-49, 'R', 51-264, 'T', 266-547 <SEE2>

A:Cross-references: GB:J05717

R:Pastuszyn, A.; Noland, B.J.; Bazan, J.F.; Fletterick, R.J.; Scallan, T.J.

J. Biol. Chem. 262, 13219-13227, 1987

A:Title: Primary sequence and structural analysis of sterol carrier protein 2 from rat

A:Reference number: A29366; MUID:88007528; PMID:3115977

A:Accession: A29366

A:Molecule type: protein

A:Residues: 425-546 <PAS>

R:Morris, H.R.; Larsen, B.S.; Billheimer, J.T.

Biochem. Biophys. Res. Commun. 154, 476-482, 1988

A:Title: A mass spectrometric study of the structure of sterol carrier protein SCP-2 from

A:Reference number: A27661; MUID:88280812; PMID:3395344

A:Accession: A27661

A:Molecule type: protein

A:Residues: 425-426, 'SV', 429-435, 'V', 437-445, 'D', 447-449, 'O', 451-487, 'N', 489-515, 'T', 517

R:Ossendorp, B.C.; van Heusden, G.P.H.; de Beer, A.L.J.; Bos, K.; Schouten, G.L.; Wirtz,

Eur. J. Biochem. 201, 233-239, 1991

A:Title: Identification of the cDNA clone which encodes the 58-kDa protein containing the

th rat peroxisomal and mitochondrial 3-oxoacyl-CoA thiolases.

A:Reference number: S17842; MUID:92007881; PMID:1915369

A:Accession: S17842

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 2-49, 'R', 51-547 <OSS>

A:Cross-references: EMBL:X60654; NID:g56771; PIDN:CAA43061.1; PID:g56773

R:Billheimer, J.T.; Strehl, L.L.; Davis, G.L.; Strauss III, J.F.; Davis, L.G.
DNA Cell Biol. 9, 159-165, 1990

A:Title: Characterization of a cDNA encoding rat sterol carrier protein-2.

A:Accession: A34584

A:Reference number: A34584; MUID:90253610; PMID:2340090

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 275-547 <BIL>

A:Cross-references: GB:M34728; NID:g206871; PIDN:AAA42120.1; PID:g206872

R:Ossendorp, B.C.; van Heusden, G.P.H.; Wirtz, K.W.A.

Biochem. Biophys. Res. Commun. 168, 631-636, 1990

A:Title: The amino acid sequence of rat liver non-specific lipid transfer protein (st

A:Reference number: A34635; MUID:90241231; PMID:2334427

A:Accession: A34635

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 33-49, 'R', 51-547 <OS2>

A:Cross-references: GB:M58287; NID:g205769; PIDN:AAA41726.1; PID:g205770

C:Keywords: alternative splicing; mitochondrion

Alignment Scores:

Pred. No.: 2.97e-42 Length: 547
Score: 49.00 Matches: 49
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.61% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-21 (1-994) x A39368 (1-547)

OY 525 CTGTTGATAGAGAGAGATATACATATGAGAGAAAGTGGTCAATAATCCTAGTGTGA 584
|||||
Db 330 LeuValAspArgGlyAspAsnThrTyrglyGlyLysTrpValIleAsnProserglygly 349
OY 585 CTGATTTCAAGGACACCCACTAGGCGCTACAGGTCCTGCTCAGTGTGCAGACTGTC 644
|||||
Db 350 LeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGlnCysAlaGluLeuCys 369
OY 645 TGGCAGCTGAGAGGGGAAGCCGGA AAA 671
|||||
Db 370 TrpGlnLeuArgGlyGlyAlaGlyLys 378

RESULT 5

S34744

sterol carrier protein - chicken

C:Species: Gallus gallus (chicken)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jul-2000

C:Accession: S34744

R:Pfeifer, S.M.; Sakuragi, N.; Ryan, A.; Johnson, A.L.; Deeley, R.G.; Billheimer, J.T

Arch. Biochem. Biophys. 304, 287-293, 1993

A:Title: Chicken sterol carrier protein 2/sterol carrier protein x: cDNA cloning reve

A:Reference number: S34744; MUID:93312016; PMID:8323294

A:Accession: S34744

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-547 <PRE>

A:Cross-references: GB:L09231; NID:g304422; PIDN:AAA02488.1; PID:g304423

Alignment Scores:

Pred. No.: 1.52e-25 Length: 547
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.51% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-21 (1-994) x S34744 (1-547)

OY 531 GATAGAGAGATATACATATGAGAGAAAGTGGTCAATAATCCTAGTGTGACTGANT 590
|||||
Db 335 AspArgGlyAspAsnThrTyrglyGlyLysTrpValIleAsnProserglyglyLeuIle 354
OY 591 TCAAGGACACCACTAGGCGCTACAGGTCCTGCTCAG 629

|||||
Db 355 SerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 367

RESULT 6

E70884

probable lipid-transfer protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: E70884

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: E70884

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-401 <COL>

A:Cross-references: GB:AL008967; GB:AL123456; NID:g3261491; PIDN:CA15585.1; PID:g262431

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: ltp1

C:Superfamily: probable 3-ketoacyl-CoA thiolase MTH1177

Alignment Scores:

Pred. No.: 3.97e-16 Length: 401
Score: 24.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.64% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x E70884 (1-401)

QY 570 AATCCTAGTGTGAGCTGATTTCAAAGGGACACCCACTAGGCGCTACAGTCTTGCTCAG 629

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 338 AsnProSerGlyGlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 357

QY 630 TGTGACAGAACTC 641

Db |||||||||||||||

Db 358 CysAlaGluLeu 361

RESULT 7

T27202

hypothetical protein Y57A10C.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C:Accession: T27202

R:Smyle, R.

submitted to the EMBL Data Library, June 1998

A:Reference number: Z20327

A:Accession: T27202

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-412 <WILL>

A:Cross-references: EMBL:AL023847; PIDN:CA19548.1; GSPDB:GN00020; CESP:Y57A10C.6

A:Experimental source: clone Y57A10C

C:Genetics:

A:Gene: CESP:Y57A10C.6

A:Map position: 2

A:Introns: 62/1; 104/1; 171/1

C:Superfamily: probable 3-ketoacyl-CoA thiolase MTH1177

Alignment Scores:

Pred. No.: 4.86e-14 Length: 412
Score: 22.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.01% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x T27202 (1-412)

QY 540 GATATACATATGAGCAAGTGGTCATAAATCCTAGTGTGACTGATTCAAGCGA 599

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 332 AspAsnThrTyGlyGlyLysTrpValIleAsnProSerGlyGlyLeuIleSerLysGly 351

QY 600 CACCCA 605

Db ||||||

Db 352 HisPro 353

RESULT 8

B90436

hypothetical protein acab-7 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001

C:Accession: B90436

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: B90436

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-387 <KUR>

A:Cross-references: GB:AE006641; NID:g13815933; PIDN:AAK42745.1; GSPDB:GN00155

C:Genetics:

A:Gene: acab-7

C:Superfamily: probable 3-ketoacyl-CoA thiolase MTH1177

Alignment Scores:

Pred. No.: 0.167 Length: 387
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.18% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x B90436 (1-387)

QY 594 AAGGACACCCACTAGGCGCTACAGGCTT 623

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 333 LysGlyHisProLeuGlyAlaThrGlyLeu 342

RESULT 9

T45470

ATP-binding ABC transport protein [imported] - Erysipelothrix rhusiopathiae

C:Species: Erysipelothrix rhusiopathiae

C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 17-Mar-2000

C:Accession: T45470

R:Antori, R.P.; Turner, B.; Eamens, G.J.; Delaney, S.F.; Chin, J.C.

submitted to the EMBL Data Library, March 1996

A:Description: Cloning and nucleotide sequencing of ewlA, the gene encoding an immuno

A:Reference number: Z22979

A:Accession: T45470

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-246 <ANID>

A:Cross-references: EMBL:U52850; PIDN:AAB01157.1

A:Experimental source: strain VRS 229

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Alignment Scores:

Pred. No.: 1.97 Length: 246
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.90% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x T45470 (1-246)

OY 782 AAGCACTGCGCGCTTCCGCAAAACC 756
|||||
Db 184 LygIuLeuAlaAlaSerGlysthr 192

RESULT 10

A96559

hypothetical protein F5F19.1 [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: A96559

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Taiton,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A96559

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-351 <STO>

A:Cross-references: GB:AE005173; NID:g4220443; PIDN:AAD12670.1; GSPDB:GN00141

C:Genetics:

A:Gene: F5F19.1

A:Map position: 1

Alignment Scores:

Pred. No.:	1.87	Length:	351
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.87%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-21 (1-994) x A96559 (1-351)

OY 150 AGTTAGATGAGTGATGCATCTTAA 176

|||||
Db 320 SerLeuAspGluValMetAlaSerLys 328

RESULT 11

G69205

lipid-transfer protein (sterol or nonspecific) - Methanobacterium thermoautotrophicum (S

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999

C:Accession: G69205

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicalre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
K1, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A:Reference number: A69000; MUID:98037514; PMID:9371463

A:Accession: G69205

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-383 <MTH>

A:Cross-references: GB:AE000857; GB:AE000666; NID:g2621876; PIDN:AAB85293.1; PID:g262188

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH793

C:Superfamily: probable 3-ketoacyl-CoA thiolase MTH1177

Alignment Scores:

Pred. No.:	1.85	Length:	383
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.87%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-21 (1-994) x G69205 (1-383)

OY 645 TGGCAGCTGAGAGGGAAGCCGAAAA 671

|||||
Db 347 TrpGluLeuArgGlyGluAlaGlyLys 355

RESULT 12

A40015

sterol carrier protein 2 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 02-Sep-2000

C:Accession: A40015

R:Moncechl, D.; Pastuszyn, A.; Scallan, T.J.

J. Biol. Chem. 266, 9885-9892, 1991

A:Title: cDNA sequence and bacterial expression of mouse liver sterol carrier protein

A:Reference number: A40015; MUID:91236770; PMID:1709640

A:Accession: A40015

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-143 <MON>

A:Cross-references: GB:M62361; NID:g200941; PIDN:AAA40099.1; PID:g200942

C:Superfamily: Caenorhabditis elegans hypothetical protein ZK892.2

C:Keywords: mitochondrion

Alignment Scores:

Pred. No.:	23.6	Length:	143
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-21 (1-994) x A40015 (1-143)

OY 760 TTCCCGAAGCCGCGAGTTCCTTT 783

|||||
Db 3 PheProGluAlaAlaSerSerPhe 10

RESULT 13

T19267

hypothetical protein C14B1.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T19267

R:Harris, B.
submitted to the EMBL Data Library, September 1994

A:Reference number: Z19099

A:Accession: T19267

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-231 <WIL>

A:Cross-references: EMBL:Z37139; PIDN:CAA85488.1; GSPDB:GN00021; CESP:C14B1.8

A:Experimental source: clone C14B1

C:Genetics:

A:Gene: CESP:C14B1.8

A:Map position: 3

A:Introns: 25/1; 58/1; 117/2; 146/2; 181/2

Alignment Scores:

Pred. No.:	22	Length:	231
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-21 (1-994) x T19267 (1-231)

OY 331 GCTGTTGAAGAAAAAGCATTA 354

|||||
Db 143 AlaArgLeuLysLysAlaLeu 150

RESULT 14

T35889
hypothetical protein SC9B10.20c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C;Accession: T35889
R;Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, November 1997
A;Reference number: Z21592
A;Accession: T35889
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-267 <OLI>
A;Cross-references: EMBL:AL009204; PIDN:CAA15810.1; GSPDB:GN00070; SCOEDB:SC9B10.20c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC9B10.20c
C;Superfamily: Mycobacterium hypothetical protein RV0911

Alignment Scores:
Pred. No.: 21.6 Length: 267
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x T35889 (1-267)

QY 468 AACGAACCTCCTTACTTATGAAGCA 491
|||||
Db 154 AsnGluleuLeuThrTyrGluAla 161

RESULT 15

S58137

gene 7 protein - phage SPPI

C;Species: phage SPPI

C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 11-May-2000

C;Accession: S58137; S24456; T42271

R;Becker, B.; Gassel, M.; Tavares, P.; Lurz, R.; Alonso, J.C.

submitted to the EMBL Data Library, July 1995

A;Description: Head morphogenesis of the Bacillus subtilis bacteriophage SPPI.

A;Reference number: S58137

A;Accession: S58137

A;Molecule type: DNA

A;Residues: 1-308 <BEC>

A;Cross-references: EMBL:X89721; NID:g1052805; PIDN:CAA61865.1; PID:g1052806

R;Chai, S.; Bravo, A.; Lueder, G.; Nedlin, A.; Trautner, T.A.; Alonso, J.C.

J. Mol. Biol. 224, 87-102, 1992

A;Title: Molecular analysis of the Bacillus subtilis bacteriophage SPPI region encompass

A;Reference number: S24450; MUID:92194332; PMID:1548711

A;Accession: S24456

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-43 <CHA>

A;Cross-references: EMBL:X56064; NID:g15464; PIDN:CAA39542.1; PID:g15471

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1990

R;Alonso, J.C.; Luder, G.; Stiege, A.C.; Chai, S.; Weise, F.; Trautner, T.A.

Gene 204, 201-212, 1997

A;Title: The complete nucleotide sequence and functional organization of Bacillus subtil

A;Reference number: Z22137; MUID:98094274; PMID:9434185

A;Accession: T42271

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-308 <ALO>

A;Cross-references: EMBL:X97918; PIDN:CAA66581.1

Alignment Scores:

Pred. No.: 21.1 Length: 308
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0

DB:

2

Gaps:

0

US-09-835-992A-21 (1-994) x S58137 (1-308)

QY 85 AAATTGATGAAAAATCATTAAC 108
|||||

Db 230 LysLeuAspGlyLysIleIleasn 237

Search completed: January 14, 2003, 17:40:45
Job time : 39.3073 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:54 ; Search time 15.349 Seconds
(without alignments)
5372.014 Million cell updates/sec

Title: US-09-835-992A-21

Perfect score: 314
Sequence: 1 ctccaccagtgtctctcag.....ccaccctgggtgtgtggat 994

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+_n2p.model -DEV-xlp
-Q-/cgn2_1/USPTO_spool/US09835992/runat_14012003_161525_4014/app_query.fasta_1.3932
-DB-SwissProt_40 -QFMT-fastan -SUFFIX-01ig.rsp -MINMATCH-0.1 -LOOPCL-0
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-oligo -TRANS-human40.cdl
-LIST-45 -DOCALIGN-200 -THR_SCOREquality -THR_MIN-1 -ALIGN-15 -MODE-LOCAL
-OUTFMT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USER-US09835992_@CGN_1_1_77_@runat_14012003_161525_4014 -NCPU-6 -ICPU-3
-NO_XLPHY -NO_MMALP -LARGEQUERY -NEG_SCORES-0 -WAIT -LONGLOG -DEV_TIMEOUT-120
-WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60 -FGAPOP-6 -FGAPEXT-7
-YGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database : SwissProt_40:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	223	71.0	547	1	NLTP_HUMAN
2	50	15.9	547	1	NLTP_MOUSE
3	49	15.6	547	1	NLTP_RAT
4	33	10.5	547	1	NLTP_CHICK
5	8	2.5	151	1	YCX2_OENHO
6	8	2.6	374	1	ADHL_APTAU
7	8	2.5	384	1	VANS_ENTFC
8	8	2.5	398	1	Y666_METJA
9	8	2.5	414	1	THIK_YARLI
10	8	2.5	426	1	CHLI_TOBAC
11	8	2.6	455	1	POLG_WMV2T
12	8	2.5	1005	1	Y456_CHLTR
13	8	2.6	1016	1	POLG_WMV2U
14	7	2.2	126	1	PAND_ECOLI
15	7	2.2	126	1	PAND_SALTY
16	7	2.2	126	1	PAND_YERPE
17	7	2.2	132	1	SRI4_ORYSA
18	7	2.2	159	1	MLE_TODPA

19	7	2.2	161	1	PIN_BPT4	P07068 bacterioph
20	7	2.2	167	1	YM20_MARPO	P38463 marchantia
21	7	2.3	169	1	PF03_SCHPO	Q10143 schizosacch
22	7	2.3	173	1	CRAM_MOUSE	P51437 mus musculu
23	7	2.2	199	1	HMG1_MOUSE	P40630 mus musculu
24	7	2.2	218	1	CD53_MOUSE	Q61451 mus musculu
25	7	2.3	233	1	GP42_RAT	P23505 rattus norv
26	7	2.3	233	1	SPX2_HUMAN	O60248 homo sapien
27	7	2.2	248	1	DSBG_ECO57	P58320 escherichia
28	7	2.2	248	1	DSBG_ECOLI	P77202 escherichia
29	7	2.2	272	1	PYRF_DEIRA	Q919C5 delnoccocus
30	7	2.3	290	1	TYSY_HSVAT	P12462 herpesvirus
31	7	2.3	305	1	PEX2_MOUSE	P55098 mus musculu
32	7	2.3	305	1	PEX2_RAT	P24392 rattus norv
33	7	2.3	316	1	YK95_MYCTU	Q10704 mycobacteri
34	7	2.2	322	1	YF46_AOUAE	O67500 aquilex aeo
35	7	2.2	330	1	OUTG_EMENT	P25416 emericeella
36	7	2.3	334	1	SRB7_CABEL	P54142 caenorhabdl
37	7	2.2	355	1	KARG_HOMGA	P14208 homarus gam
38	7	2.2	356	1	KARG_SCHAM	P91798 schistocerc
39	7	2.2	367	1	LHX8_MOUSE	O35652 mus musculu
40	7	2.3	389	1	G68A_DROME	O9vtn0 drosophila
41	7	2.2	399	1	CHS2_HORVU	O96562 hordeum vul
42	7	2.2	421	1	PSMR_METAC	O8t188 methanosarc
43	7	2.2	427	1	TIG_STRPN	O97sg9 streptococc
44	7	2.2	449	1	MURF_RICPR	O05953 rickettsia
45	7	2.2	450	1	MEI4_YEAST	P29467 saccharomyc

ALIGNMENTS

RESULT 1
ID NLTP_HUMAN STANDARD: PRT: 547 AA.
AC P22307; Q16622; Q15432; Q99430;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nonspecific lipid-transfer protein, mitochondrial precursor (NSL-TP)
DE (Sterol carrier protein 2) (SCP-2) (Sterol carrier protein X) (SCP-X)
DE (SCPX).
GN SCP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95213031; PubMed=7698762;
RA Ohba T., Rennett H., Pfeiffer S.M., He Z., Yamamoto R., Holt J.A.,
RA Billheimer J.T., Strauss J.F. III;
RT "The structure of the human sterol carrier protein X/sterol carrier
RT protein 2 gene (SCP2).";
RL Genomics 24:370-374(1994).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92029618; PubMed=1718316;
RA He Z., Yamamoto R., Futh E.E., Schantz L.J., Naylor S.L., George H.,
RA Billheimer J.T., Strauss J.F. III;
RT "CDNAs encoding members of a family of proteins related to human
RT sterol carrier protein 2 and assignment of the gene to human
RT chromosome 1 p21-pter.";
RL DNA Cell Biol. 10:559-569(1991).
[3]
RP SEQUENCE OF 405-547 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=9110550; PubMed=1703300;
RA Yamamoto R., Kallen C.B., Babalola G.O., Rennett H., Billheimer J.T.,
RA Strauss J.F. III;
RT "Cloning and expression of a cDNA encoding human sterol carrier
RT protein 2.";

RL Proc. Natl. Acad. Sci. U.S.A. 88:463-467(1991).
RN [4]
RP SEQUENCE OF 405-547 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE-93131254; PubMed-1483685;
RA Yamamoto R.;
RT "Localization of human sterol carrier protein 2 gene and cDNA
RT expression in COS-7 cell.";
RL Hokkaido Igaku Zasshi 67:839-848(1992).
RN [5]
RP STRUCTURE BY NMR OF SCP2.
RX MEDLINE-94063072; PubMed-8243660;
RA Szyperski T., Scheek S., Johansson J., Assmann G., Seedorf U.,
RA Wuehrich K.;
RT "NMR determination of the secondary structure and the
RT three-dimensional polypeptide backbone fold of the human sterol
RT carrier protein 2.";
RL FEBS Lett. 335:18-26(1993).
CC -1- FUNCTION: MEDIATES IN VITRO THE TRANSFER OF ALL COMMON
CC PHOSPHOLIPIDS, CHOLESTEROL AND GANGLIOSIDES BETWEEN MEMBRANES. MAY
CC PLAY A ROLE IN REGULATING STEROIDOGENESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN THE LIVER AND ALSO ASSOCIATED
CC WITH MITOCHONDRIA ESPECIALLY IN STEROIDOGENIC TISSUES. SCP-X
CC RESIDES IN THE PEROXISOME.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; SCPx (shown here) and SCP2; are
CC produced by alternative initiation.
CC -1- TISSUE SPECIFICITY: LIVER, FIBROBLASTS, AND PLACENTA.
CC -1- DISEASE: SCP2 IS PRESENT IN LOW LEVELS IN SUBJECTS WITH ZELLWEGER
CC SYNDROME (CEREBRO-HEPATIC-RENAL SYNDROME), WHOSE CELLS ARE
CC DEFICIENT IN PEROXISOMES AND WHO HAVE AN ASSOCIATED IMPAIRMENT IN
CC PLASMALOGEN AND BILE ACID SYNTHESIS AND CATABOLISM OF PHYTANIC
CC ACID AND VERY-LONG-CHAIN FATTY ACIDS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE THIOLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U11313; AAB41286.1; .
DR EMBL; U11297; AAB41286.1; JOINED.
DR EMBL; U11299; AAB41286.1; JOINED.
DR EMBL; U11300; AAB41286.1; JOINED.
DR EMBL; U11301; AAB41286.1; JOINED.
DR EMBL; U11302; AAB41286.1; JOINED.
DR EMBL; U11303; AAB41286.1; JOINED.
DR EMBL; U11304; AAB41286.1; JOINED.
DR EMBL; U11305; AAB41286.1; JOINED.
DR EMBL; U11306; AAB41286.1; JOINED.
DR EMBL; U11307; AAB41286.1; JOINED.
DR EMBL; U11308; AAB41286.1; JOINED.
DR EMBL; U11309; AAB41286.1; JOINED.
DR EMBL; U11310; AAB41286.1; JOINED.
DR EMBL; U11311; AAB41286.1; JOINED.
DR EMBL; U11312; AAB41286.1; JOINED.
DR EMBL; M75883; AAA03557.1; .
DR EMBL; M75884; AAA03558.1; ALT_INIT.
DR EMBL; M55421; AAA03559.1; ALT_SEQ.
DR EMBL; S52450; AAB24921.1; .
DR PIR; A39010; A39010.
DR HSSP; P27796; 1PXT.
DR Genew; HGNC:10606; SCP2.
DR MIM; 184755; .
DR InterPro; IPR003033; SCP2.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02036; SCP2; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLASE_1; 1.

DR PROSITE; PS00737; THIOLASE_2; 1.
DR PROSITE; PS00099; THIOLASE_3; FALSE_NEG.
DR PROSITE; PS00342; MICROBODIES_CTER; 1.
KW Lipid-binding; Transport; Mitochondrion; Peroxisome; Transit peptide;
KW Alternative initiation.
FT CHAIN 1 547
FT TRANSIT 405 424
FT CHAIN 425 547
FT INIT_MET 405 405
FT ACT_SITE 94 94
FT SITE 494 494
FT SITE 545 547
FT CONFLICT 10 10
FT CONFLICT 393 393
FT CONFLICT 472 472
FT CONFLICT 482 482
FT CONFLICT 501 501
FT CONFLICT 522 522
60 SEQUENCE 547 AA; 58993 MW; 29F7551465C7143A CRC64;

Alignment Scores:
Pred. No.: 3.06e-224 Length: 547
Score: 223.00 Matches: 223
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 71.02% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-21 (1-994) x NLRP_HUMAN (1-547)

QY 3 CACCAGTGTGCTCTCAGATGTTGGGTATGCTGGAAGAACATATGGAATAATGGA 62
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 156 HisProvalAlaProGlnMetPheGlyTyrAlaGlyLysGlnHisMetGluLysTyrGly 175
QY 63 ACAAAATGGAACACTTGCAAAATTTGGATGGAATAATCATAAACATTCAGTTAATAAC 122
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 176 ThrLysIleGlnHisPheAlaLysIleGlyTyrLysAsnHisLysHisSerValAsn 195
QY 123 CCGTATTCACAGTTCACAGATGAATACAGTTAGATGAAGTGCATGCAATGAAGAATT 182
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 196 ProTyrSerGlnPheGlnAspGluTyrSerLeuAspGluValMetAlaSerLysGluVal 215
QY 183 TTGATTTTGTGACTATCTTACAATGTTGCCACTTCAGATGGTGCAGACCAATT 242
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216 PheAspPheLeuThrIleLeuGlnCysCysProThrSerAspGlyAlaAlaAlaIle 235
QY 243 TTGGCCAGTGAAGCATTTGTACAGAGTATGCCCTCAATCCAAAGCTGTGAAATTTTG 302
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 236 LeuAlaSerGlnAlaPheValGlnLysTyrGlyLeuGlnSerLysAlaValGluIleLeu 255
QY 303 GCACAAGAAATGATGACTGATTGGCAAGCTGCTTTGAAGAAAAAAGCATTAATAATG 362
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 256 AlaGlnGlnMetMetThrAspLeuProSerSerPheGlnGluLysSerIleIleLysMet 275
QY 363 GTTGCTTTGATATGAGTAAGAAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGACA 422
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 276 ValGlyPheAspMetSerLysGlnAlaAlaArgLysCysTyrGluLysSerGlyLeuThr 295
QY 423 CCAATGATATTGACGTAATAGAACTTCACGATGCTTTTCTACCAACGAACTCTTACT 482
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 296 ProAsnAspIleAspValIleGluLeuHisAspCysPheSerThrAsnGluLeuLeuThr 315
QY 483 TATGAAGCAGCTGGGACTCTGTCCAGAAGGACAAAGCTGCTGTGATAGAGAGAT 542
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 316 TyrGluAlaLeuGlyLeuGlyCysProGlnGlyGlnGlyAlaThrLeuValAspArgGlyAsp 335
QY 543 AATACATATGAGAGAAAGTGGGTATTAATCTAGTGGTACATGATTTCAAGGACAC 602
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 336 AsnThrTyrGlyGlyLysTyrValIleAsnProSerGlyGlyLeuIleSerLysGlnHis 355
QY 603 CCACTAGGCGGTACAGGTCTTGCTCAGTGTGCAAGAACTGTGCTGGCAGCTGAGAGGGA 662

|||||
Db 356 ProLeuGlyAlaThrGlyLeuAlaGlnCysAlaGluLeuLeuArgGlyGlu 375
OY 663 CCCCCGAAAA 671
Db 376 AlaGlyLys 378
RESULT 2
NLTP_MOUSE
ID NLTP_MOUSE STANDARD; PRT; 547 AA.
AC P32020;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nonspecific lipid-transfer protein, mitochondrial precursor (NSL-Tp)
DE (Sterol carrier protein 2) (SCP-2) (Sterol carrier protein X) (SCP-X)
DE (SCPX).
GN SCP2 OR SCP-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93154580; PubMed=8428655;
RA Seedorf U., Raabe M., Assmann G.;
RT "Cloning, expression and sequences of mouse sterol-carrier protein-x-
RT encoding cDNAs and a related pseudogene.";
RL Gene 123:165-172(1993).
RN [2]
RP SEQUENCE OF 405-547 FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=91236770; PubMed=1709640;
RA Moncechl D.T., Pastuszyn A., Scallan T.J.;
RT "cDNA sequence and bacterial expression of mouse liver sterol carrier
RT protein-2.";
RL J. Biol. Chem. 266:9885-9892(1991).
RN [3]
RP SEQUENCE OF 1-23 FROM N.A.
RC STRAIN=129/SV; TISSUE=Liver;
RX MEDLINE=96354847; PubMed=8751375;
RA Raabe M., Seedorf U., Hamelster H., Ellinghaus P., Assmann G.;
RT "Structure and chromosomal assignment of the murine sterol carrier
RT protein 2 gene (Scp2) and two related pseudogenes by in situ
RT hybridization.";
RL Cytogenet. Cell Genet. 73:279-281(1996).
CC -1- FUNCTION: MEDIATES IN VITRO THE TRANSFER OF ALL COMMON
CC PHOSPHOLIPIDS, CHOLESTEROL AND GLYCOSIDES BETWEEN MEMBRANES. MAY
CC PLAY A ROLE IN REGULATING STEROIDOGENESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN THE LIVER AND ALSO ASSOCIATED
CC WITH MITOCHONDRIA ESPECIALLY IN STEROIDogenic TISSUES. SCP-X
CC RESIDES IN THE PEROXISOME.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; SCPx (shown here) and SCP2; are
CC produced by alternative initiation.
CC -1- TISSUE SPECIFICITY: PRESENT AT LOW LEVELS IN ALL TISSUES
CC EXAMINED BUT EXPRESSED PREDOMINANTLY IN THE LIVER.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE THIOLEASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL: M91458; AAA40098.1; -
CC EMBL: M62361; AAA40099.1; -
CC EMBL: X91150; CAA62592.1; -
CC PIR: A40015; A40015.
CC PIR: JU0157; JU0157.

DR MGD; MGI:98254; SCP2.
DR InterPro; IPR003033; SCP2.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02036; SCP2; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLEASE_1; 1.
DR PROSITE; PS00737; THIOLEASE_2; 1.
DR PROSITE; PS00099; THIOLEASE_3; 1.
DR PROSITE; PS00342; MICROBODIES_CTER; 1.
KW Lipid-binding; Transport; Mitochondrion; Peroxisome; Transit peptide;
KW Alternative initiation.
FT CHAIN 1 547
FT TRANSIT 405 424
FT CHAIN 425 547
FT INIT_MET 405 405
FT ACT_SITE 94 94
FT SITE 494 494
FT SITE 545 547
FT CONFLICT 478 478
SO SEQUENCE 547 AA; 59158 MW; BAB02F986D16B18E CRC64;
Alignment Scores:
Pred. No.: 1.26e-43 Length: 547
Score: 50.00 Matches: 50
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.92% Indels: 0
DB: 1 Gaps: 0
US-09-835-992a-21 (1-994) x NLTP_MOUSE (1-547)
OY 522 ACGCTGTTGATAGAGAGATTAATCATATGAGAAAGTGGTCATAATCCTAGTGGT 581
Db 329 ThrLeuValAspArgGlyAspAsnThrTyrGlyGlyLysTyrValIleAsnProserGly 348
OY 582 GGACGTGATTTCAAAGGACACCCACTAGCGCTTCTTCAGTGTGACAGACTC 641
Db 349 GlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGlnCysAlaGluLeu 368
OY 642 TGCTGGCAGCTGAGAGGGGAGCCGAAAA 671
Db 369 CysTirpGlnLeuArgGlyGluAlaGlyLys 378
RESULT 3
NLTP_RAT
ID NLTP_RAT STANDARD; PRT; 547 AA.
AC P11915; Q63383;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nonspecific lipid-transfer protein, mitochondrial precursor (NSL-Tp)
DE (Sterol carrier protein 2) (SCP-2) (Sterol carrier protein X) (SCP-X)
DE (SCPX).
GN SCP2 OR SCP-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=91093192; PubMed=1985920;
RA Seedorf U., Assmann G.;
RT "Cloning, expression, and nucleotide sequence of rat liver sterol
RT carrier protein 2 cDNAs.";
RL J. Biol. Chem. 266:630-636(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91239563; PubMed=2034675;

RA Mori T., Tsukamoto T., Mori H., Tashiro Y., Fujiki Y.;
RT "Molecular cloning and deduced amino acid sequence of nonspecific
RT lipid transfer protein (sterol carrier protein 2) of rat liver: a
RT higher molecular mass (60 kDa) protein contains the primary sequence
RT of nonspecific lipid transfer protein as its C-terminal part.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:4338-4342(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92007881; PubMed=1915369;
RA Ossendorp B.C., van Heusden G.P.H., de Beer A.L.J., Bos K.,
RA Schouten G.L., Wirtz K.W.A.;
RT "Identification of the cDNA clone which encodes the 58-kDa protein
RT containing the amino acid sequence of rat liver non-specific lipid-
RT transfer protein (sterol-carrier protein 2). Homology with rat
RT peroxisomal and mitochondrial 3-oxoacyl-CoA thiolases.";
RL Eur. J. Biochem. 201:233-239(1991).
RN [4]
RP SEQUENCE OF 33-547 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90241231; PubMed=2334427;
RA Ossendorp B.C., van Heusden G.P.H., Wirtz K.W.A.;
RT "The amino acid sequence of rat liver non-specific lipid transfer
RT protein (sterol carrier protein 2) is present in a high molecular
RT weight protein: evidence from cDNA analysis.";
RL Biochem. Biophys. Res. Commun. 168:631-636(1990).
RN [5]
RP SEQUENCE OF 275-547 FROM N.A.
RC STRAIN=CD Charles River; TISSUE=Liver;
RX MEDLINE=90253610; PubMed=2340090;
RA Billheimer J.T., Strehl L.L., Davis G.L., Strauss J.F. III,
RA Davis L.G.;
RT "Characterization of a cDNA encoding rat sterol carrier protein2.";
RL DNA Cell Biol. 9:159-165(1990).
RN [6]
RP SEQUENCE OF 425-547.
RC TISSUE=Liver;
RX MEDLINE=88007528; PubMed=3115977;
RA Pastuszyn A., Noland B.J., Bazan J.F., Fletterick R.J., Scallan T.J.;
RT "Primary sequence and structural analysis of sterol carrier protein 2
RT from rat liver: homology with immunoglobulins.";
RL J. Biol. Chem. 262:13219-13227(1987).
RN [7]
RP SEQUENCE OF 425-547.
RC TISSUE=Liver;
RX MEDLINE=88280812; PubMed=3395344;
RA Morris H.R., Larsen B.S., Billheimer J.T.;
RT "A mass spectrometric study of the structure of sterol carrier
RT protein SCP2 from rat liver.";
RL Biochem. Biophys. Res. Commun. 154:476-482(1988).
CC -1- FUNCTION: MEDIATES IN VITRO THE TRANSFER OF ALL COMMON
CC PHOSPHOLIPIDS, CHOLESTEROL AND GANGLIOSIDES BETWEEN MEMBRANES. MAY
CC PLAY A ROLE IN REGULATING STEROIDOGENESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN THE LIVER AND ALSO ASSOCIATED
CC WITH MITOCHONDRIA ESPECIALLY IN STEROIDOGENIC TISSUES. SCP-X
CC RESIDES IN THE PEROXISOME.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; SCPx (shown here) and SCP2; are
CC produced by alternative initiation.
CC -1- TISSUE SPECIFICITY: LIVER > INTESTINE > BRAIN > LUNG, COLON,
CC STOMACH, SPLEEN, KIDNEY, HEART, AND OVARY.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE THIOLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62763; AAAA40622.1; -;
DR EMBL; M62763; AAAA40623.1; ALT_INIT.
DR EMBL; M57454; AAAA2121.1; -;

DR EMBL; M57453; AAAA2122.1; -;
DR EMBL; M58287; AAAA1726.1; -;
DR EMBL; M34728; AAAA2120.1; -;
DR EMBL; X60654; CAA43060.1; ALT_INIT.
DR EMBL; X60654; CAA43061.1; -;
DR PIR; A27661; A27661.
DR PIR; A29366; A29366.
DR PIR; A34584; A34584.
DR PIR; A34635; A34635.
DR PIR; A39054; A39054.
DR InterPro; IPR003033; SCP2.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF02036; SCP2; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00099; THIOLASE_3; FALSE_NEG.
DR PROSITE; PS00098; MICROBODIES_CTER; 1.
DR PROSITE; PS00342; MICROBODIES_CTER; 1.
DR PROSITE; PS00737; THIOLASE_2; 1.
KW Lipid-binding; Transport; Mitochondrion; Peroxisome; Transit peptide;
KW Alternative initiation.
FT CHAIN 1 547
FT TRANSIT 405 424
FT CHAIN 425 547
FT INIT_MET 405 405
FT ACT_SITE 94 94
FT SITE 494 494
FT SITE 545 547
FT CONFLICT 12 12
FT CONFLICT 50 50
FT CONFLICT 265 265
FT CONFLICT 427 428
FT CONFLICT 436 436
FT CONFLICT 446 446
FT CONFLICT 450 450
FT CONFLICT 488 488
FT CONFLICT 516 516
FT CONFLICT 526 526
FT CONFLICT 537 537
FT CONFLICT 543 543
SQ SEQUENCE 547 AA; 58813 MW; DDD1B435D2DC6AFB CRC64;

Alignment Scores:
Pred. No.: 1.39e-42 Length: 547
Score: 49.00 Matches: 49
Percent Similarity: 100.00% Conserves: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.61% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-21 (1-994) x NLRP_RAT (1-547)
QY 525 CTGCTGATAGAGAGAGATATACATATGAGGAAGTGGCTCATTAATCCTAGTGTGCA 584
Db 330 LeuValAspArgGlyAspAsnThrTyrGlyGlyLysTrpValIleAsnProSerGlyGly 349
QY 585 CTGATTTCAAAGGAGACACCCACTAGCGGCTACAGGTCCTGCTCAGTGTGCAACTCTGC 644
Db 350 LeuIleSerIysGlyHisProLeuGlyAlaThrGlyLeuAlaGlnCysAlaGluLeuCys 369
QY 645 TGGCAGCTGAGAGGGGAGGCCGGAAGA 671
Db 370 TrpGlnLeuArgGlyGlyAlaGlyLys 378
RESULT 4
NLRP_CHICK-
ID NLRP_CHICK STANDARD; PRT; 547 AA.
AC Q07598;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Nonspecific lipid-transfer protein, mitochondrial precursor (NSL-TP)
DE (Sterol carrier protein 2) (SCP-2) (Sterol carrier protein X) (SCP-X)
DE (SCPX) (Fragment).
GN SCP2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID-9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-white leghorn; TISSUE-Liver;
RX MEDLINE-93312016; PubMed-8323294;
RA Pfeiffer S.M., Sakurai N., Ryan A., Johnson A.L., Deeley R.G.,
RA Bilheimer J.T., Baker M.E., Straus J.F. III;
RT *Chicken sterol carrier protein 2/sterol carrier protein x: cDNA
RT cloning reveals evolutionary conservation of structure and regulated
RT expression.";
RL Arch. Biochem. Biophys. 304:287-293(1993).
CC -1- FUNCTION: MEDIATES IN VITRO THE TRANSFER OF ALL COMMON
CC PHOSPHOLIPIDS, CHOLESTEROL AND GLYCOSIDES BETWEEN MEMBRANES. MAY
CC PLAY A ROLE IN REGULATING STEROIDOGENESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN THE LIVER AND ALSO ASSOCIATED
CC WITH MITOCHONDRIA ESPECIALLY IN STEROIDGENIC TISSUES. SCP-X
CC RESIDES IN THE PEROXISOME.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; SCPx (shown here) and SCP2; are
CC produced by alternative initiation.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE LIVER,
CC INTESTINE AND OVARIAN GRANULOSA CELLS.
CC -1- DEVELOPMENTAL STAGE: SCP-X LEVELS REMAIN UNCHANGED DURING DAY 20
CC EMBRYO TO 4 WEEKS POSTHATCH. A 10-FOLD INCREASE IN SCP-2 LEVEL IS
CC SEEN BY 1 WEEK POSTHATCH AND DECLINES SLIGHTLY BETWEEN 3 AND 4
CC WEEKS POSTHATCH.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE THIOLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L09231; AAA02488.1; -
DR InterPro: IPR003033; SCP2.
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF00108; thiolase; 1.
DR Pfam: PF02803; SCP2; 1.
DR Pfam: PF02803; thiolase_C; 1.
DR PROSITE: PS00098; THIOLASE_1; 1.
DR PROSITE: PS00737; THIOLASE_2; 1.
DR PROSITE: PS00099; THIOLASE_3; FALSE_NEG.
DR PROSITE: PS00342; MICROBODIES_CTER; 1.
KW Lipid-binding; Transport; Mitochondrion; Peroxisome; Transit peptide;
KW Alternative initiation.
FT CHAIN <1 547
FT NONSPECIFIC LIPID-TRANSFER PROTEIN, SCPX
FT ISOFORM.
FT MITOCHONDRION (BY SIMILARITY).
FT CHAIN 428 547
FT NONSPECIFIC LIPID-TRANSFER PROTEIN, SCP2
FT ISOFORM (BY SIMILARITY).
FT INIT_MET 408 408
FT ACT_SITE 97 97
FT SITE 494 494
FT ESSENTIAL FOR TRANSPORT OF LIPIDS (BY
FT SIMILARITY).
FT SITE 545 547
FT MICROBODY TARGETING SIGNAL (POTENTIAL).
SO SEQUENCE 547 AA; 58691 MW; CAFCDDE4A1CB3B8 CRC64;

Alignment Scores:
Pred. No.: 7.06e-26
Score: 33.00
Percent Similarity: 100.00%

Length: 547
Matches: 33
Conservative: 0

Best Local Similarity: 100.00%
Query Match: 10.51%
DB: 1
Mismatches: 0
Indels: 0
Gaps: 0
US-09-835-992a-21 (1-994) x YCX2_OENHO (1-151)
QY 531 GATAGAGGATATATACATATGAGGAAGTGGTCATAAATCTAGTGTGACTGATT 590
Db 335 Asparagilyasparanthrttygylglylstrvalleasnproserglyglyleulle 354
QY 591 TCAAGGACACCCACTAGGCGCTACAGCTTGTCTCAG 629
Db 355 Serlysglyhisproleuglyalathrglyleualagln 367
RESULT 5
YCX2_OENHO
ID YCX2_OENHO STANDARD; PRT; 151 AA.
AC O9MTP2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 17.5 kDa protein in trnf-trnl intergenic region (ORF151).
OS Oenothera hookeri (Hooker's evening primrose).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Myrtales; Onagraceae; Oenothera.
OX NCBI_TaxID-85636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Johansen;
RX MEDLINE-20309318; PubMed-10852478;
RA Hupfer H., Swiatek M., Hornung S., Herrmann R.G., Maier R.M.,
RA Chiu W.-L., Sears B.;
RT "Complete nucleotide sequence of the Oenothera elata plastid
RT chromosome, representing plastome I of the five distinguishable
RT Euenothera plastomes.";
RL Mol. Gen. Genet. 263:581-585(2000).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ271079; CAB67133.1; -
KW Chloroplast; Hypothetical protein.
SO SEQUENCE 151 AA; 17459 MW; 2DB3FEB3584BE703 CRC64;

Alignment Scores:
Pred. No.: 11.1
Score: 8.00
Length: 151
Matches: 8
Conservative: 0
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Mismatches: 0
Query Match: 2.55%
Indels: 0
DB: 1
Gaps: 0
US-09-835-992a-21 (1-994) x YCX2_OENHO (1-151)
QY 619 GTCCTGCTCAGTGTGACAGACTCT 642
Db 41 ValLeuLeuSerValGlnAsnSer 48
RESULT 6
ADH1_APTAU
ID ADH1_APTAU STANDARD; PRT; 374 AA.
AC P49645;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alcohol dehydrogenase I (EC 1.1.1.1).

```
OS Apterix australis (Brown kiwi).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Palaeognathae; Apterygiformes; Apterygidae;
OC Apterix.
OX NCBI_TaxID=8822;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95331397; PubMed=7541757;
RA Hjelmqvist L., Metsis M., Persson H., Hoeoeg J.-O., McLennan J.,
RA Joernvall H.;
RT "Alcohol dehydrogenase of class I: kiwi liver enzyme, parallel
RT evolution in separate vertebrate lines, and correlation with 12S rRNA
RT patterns.";
RT FEBS Lett. 367:306-310(1995).
CC -1- CATALYTIC ACTIVITY: An alcohol + NAD(+) = an aldehyde or ketone +
CC NADH.
CC -1- COFACTOR: ZINC.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; S78778; AAC60755.2; -.
DR HSSP; P00327; 2OHX.
DR InterPro; IPR002328; ADH_zinc.
DR InterPro; IPR002085; Adh_zn_family.
DR Pfam; PF00107; adh_zinc; 1.
DR PROSITE; PS00059; ADH_ZINC; 1.
KW Oxidoreductase; Zinc; NAD; Acetylation.
FT INIT_MET 0
FT MOD_RES 0
FT METAL 1
FT METAL 46
FT METAL 67
FT METAL 97
FT METAL 97
FT METAL 100
FT METAL 103
FT METAL 111
FT METAL 174
SQ SEQUENCE 374 AA; 39500 MW; 4D1BD387891F63EF CRC64;

Alignment Scores:
Pred. No.: 9.49
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.58%
DB: 1
Gaps: 0

US-09-835-992a-21 (1-994) x ADH1_APTAU (1-374)
QY 247 GCCAAATGCTGCTGCAGCACCA 224
Db 158 Alalysilealalalalalapro 165

RESULT 7
VANS_ENTFC
ID VANS_ENTFC STANDARD; PRT; 384 AA.
AC Q06240;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sensor protein vans (EC 2.7.3.-) (Vancomycin resistance protein vans)
DE (Vancomycin histidine protein kinase).
GN VANS.
OS Enterococcus faecium (Streptococcus faecium).
OG Plasmid pIP816.
```

```
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1352;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BM4147; TRANSPOSON=Tn1546;
RX MEDLINE=93106944; PubMed=8380148;
RA Arthur M., Molinas C., Depardieu F., Courvalin P.;
RT "Characterization of Tn1546, a Tn3-related transposon conferring
RT glycopeptide resistance by synthesis of depsipeptide peptidoglycan
RT precursors in Enterococcus faecium BM4147.";
RL J. Bacteriol. 175:117-127(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BM4147;
RX MEDLINE=92210502; PubMed=1556077;
RA Arthur M., Molinas C., Courvalin P.;
RT "The Vans-Vanr two-component regulatory system controls synthesis of
RT depsipeptide peptidoglycan precursors in Enterococcus faecium
RT BM4147.";
RL J. Bacteriol. 174:2582-2591(1992).
CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM VANS/VANR.
CC ACTIVATES THE TRANSCRIPTION OF VANH, VANA AND VANX IN RESPONSE TO
CC VANCOMYCIN WHICH RESULTS IN VANCOMYCIN RESISTANCE. VANS ACTIVATES
CC VANR BY PHOSPHORYLATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M97297; AAA65954.1; -.
DR EMBL; M68910; AAA24788.1; -.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004359; HIS_KIN_sig.
DR InterPro; IPR003661; His_kinA.
DR Pfam; PF00512; signal; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HSKA; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Sensory transduction; Transferase; Kinase; Transmembrane; Cell wall;
KW Antibiotic resistance; Phosphorylation; Plasmid.
FT TRANSMEM 21
FT TRANSMEM 41
FT TRANSMEM 76
FT DOMAIN 161
FT MOD_RES 164
FT MOD_RES 164
SQ SEQUENCE 384 AA; 43915 MW; D63099844F37C35A CRC64;

Alignment Scores:
Pred. No.: 9.45
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.55%
DB: 1
Gaps: 0

US-09-835-992a-21 (1-994) x VANS_ENTFC (1-384)
QY 58 ATGACAAATGACACTTTG 81
Db 130 MetGluGlnLysLeuAsnThrLeu 137

RESULT 8
V666_METJA
ID V666_METJA STANDARD; PRT; 398 AA.
AC Q58080;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
```


DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative molybdopterin biosynthesis protein MJ0666.
CN MJ0666.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weldman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadov P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- PATHWAY: MOLYBDENUM COFACTOR BIOSYNTHESIS (BY SIMILARITY).
CC -1- SIMILARITY: STRONG, TO BACTERIAL MOEA PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67514; AAB98661.1; .
DR TIGR; MJ0666; .
DR InterPro; IPR001453; MOCF_biosynth.
DR InterPro; IPR005111; MoeA_C.
DR InterPro; IPR005110; MoeA_N.
DR Pfam; PF009994; MOCF_biosynth; 1.
DR Pfam; PF03453; MoeA_N; 1.
DR Pfam; PF03454; MoeA_C; 1.
DR PRODOM; PD002460; MOCF_biosynth; 1.
DR TIGRFAMS; TIGR00177; molyb-syn; 1.
DR PROSITE; PS01079; MOCF_BIOSYNTHESIS_2; 1.
KW Hypothetical protein; Molybdenum cofactor biosynthesis;
KW Complete proteome.
SQ SEQUENCE 398 AA; 44418 MW; 81747A9908607ACB CRC64;

Alignment Scores:
Pred. No.: 9.39 Length: 398
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-21 (1-994) x Y666_METUA (1-398)
QY 85 AAATTGATGCAAAATCATTAAC 108
DB 204 LysleuAspGlyLysIleasn 211

RESULT 9
THIK_YARLI
ID THIK_YARLI STANDARD; PRT; 414 AA.
AC Q05493;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 3-ketoacyl-CoA thiolase, peroxisomal precursor (EC 2.3.1.16) (Beta-
DE ketothiolase) (Acetyl-CoA acyltransferase) (Peroxisomal 3-oxoacyl-
DE CoA thiolase).
GN POT1.

OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93387313; PubMed=7916689;
RA Berninger G., Schmidtchen R., Casel G., Knoerr A.,
RA Rautenstrauss K., Kunau W.-H., Schweizer E.;
RT "Structure and metabolic control of the Yarrowia lipolytica
RT peroxisomal 3-oxoacyl-CoA-thiolase gene.";
RL Eur. J. Biochem. 216:607-613(1993).
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + acetyl-CoA = CoA + 3-oxoacyl-CoA.
CC -1- PATHWAY: ONE OF THE ENZYMES OF THE FATTY ACID BETA-OXIDATION
CC CYCLE IN PEROXISOMES.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Peroxisomal.
CC -1- SIMILARITY: BELONGS TO THE THIOLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69988; CAA49605.1; .
DR PIR; S31331; S31331.
DR PIR; S36838; S36838.
DR HSSP; P27796; 1PXT.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLASE_1; 1.
DR PROSITE; PS00099; THIOLASE_3; 1.
DR PROSITE; PS00737; THIOLASE_2; 1.
KW Fatty acid metabolism; Transferase; Acyltransferase; Peroxisome;
KW Transit peptide.
FT TRANSIT 1 ? MICROBODY.
FT CHAIN ? 414 3-KETOACYL-COA THIOLASE.
FT ACT_SITE 115 115 SUBSTRATE BINDING (BY SIMILARITY).
FT ACT_SITE 400 400 BASE (BY SIMILARITY).
SQ SEQUENCE 414 AA; 43059 MW; 29D5F83A8FC792B1 CRC64;

Alignment Scores:
Pred. No.: 9.33 Length: 414
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-21 (1-994) x THIK_YARLI (1-414)
QY 597 GGACACCACTAGCGCTACAGT 620
DB 369 GlyHisProLeuGlyAlaThrGly 376

RESULT 10
CHLI_TOBAC
ID CHLI_TOBAC STANDARD; PRT; 426 AA.
AC O22436;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Magnesium-chelataase subunit chli, chloroplast precursor (Mg-
DE protoporphyrin IX chelataase).
GN CHLI.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.


```

OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. SRL;
RX MEDLINE=98088034; PubMed=9426628;
RA Kruse E., Mock H.-P., Grimm B.;
RT *Isolation and characterisation of tobacco (Nicotiana tabacum) cDNA
RT clones encoding proteins involved in magnesium chelation into
RT protoporphyrin IX.";
RL Plant Mol. Biol. 35:1053-1056(1997).
CC -I- FUNCTION: Involved in chlorophyll biosynthesis; introduces a
CC magnesium ion into protoporphyrin IX to yield Mg-protoporphyrin
CC IX.
CC -I- PATHWAY: Chlorophyll biosynthesis.
CC -I- SUBCELLULAR LOCATION: Chloroplast stroma.
CC -I- TISSUE SPECIFICITY: Strongly expressed in young leaves and to a
CC lesser extent in mature leaves.
CC -I- SIMILARITY: BELONGS TO THE MG-CHELATASE SUBUNITS D/I FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF014053; AAB97153.1;-
CC InterPro; IPR003593; AAA_Atpase.
CC InterPro; IPR000523; Mg_chelatase_ch11.
CC Pfam; PF01078; Mg_chelatase; 1.
CC SMART; SM00382; AAA; 1.
CC Photosynthesis; Chlorophyll biosynthesis; Chloroplast;
CC Transit peptide; ATP-binding.
CC TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
CC CHAIN ? 426 MAGNESIUM-CHELATASE SUBUNIT CHL1.
CC NP_BIND 121 128 ATP (POTENTIAL).
CC SEQUENCE 426 AA; 46627 MW; 42B7A6F04E1C3274 CRC64;

Alignment Scores:
Pred. No.: 9.28 Length: 426
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
Gaps: 0

US-09-835-992A-21 (1-994) x CHLI_TOBAC (1-426)
QY 228 GCTGCAGCAGCAATTTGGCCAGT 251
Db 11 AlaAlaAlaAlaIleuAlaSer 18

RESULT 11
POLG_WMV2T STANDARD; PRT; 455 AA.
AC 089251;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Genome polyprotein [contains: Nuclear inclusion protein B (NI-B) (NIB)
DE (RNA-directed RNA polymerase) (EC 2.7.7.48); Coat protein (CP)]
DE (Fragment).
OS Watermelon mosaic virus II (isolate Tonga) (Vanilla necrosis
OS polyvirus) (VNV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=148359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93228466; PubMed=8470960;
RA Wang Y.Y., Beck D.L., Gardner R.C., Pearson M.N.;
RT *Nucleotide sequence, serology and symptomology suggest that vanilla

```

```

RT necrosis polyvirus is a strain of watermelon mosaic virus II.";
RL Arch. Virol. 129:93-103(1993).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -1- PFM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
CC POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC INDIVIDUAL PROTEINS.
CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L22907; AAA48497.2; ALT_INIT.
CC DR InterPro: IPR001592; Poly_coat.
CC DR InterPro: IPR001205; RNA_pol_P3D.
CC DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC DR Pfam: PF00767; Poly_coat; 1.
CC KW Transferase; RNA-directed RNA polymerase; Coat protein; Polyprotein.
CC FT NON_TER 1 1
CC FT CHAIN <1 172 NUCLEAR INCLUSION PROTEIN B.
CC FT CHAIN 173 455 COAT PROTEIN.
CC FT SEQUENCE 455 AA; 51607 MW; 937AFDEAE9A52EC1 CRC64;
CC -----
Alignment Scores:
Pred. No.: 9.18 Length: 455
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.58% Indels: 0
DB: 1 Gaps: 0
-----
US-09-835-992A-21 (1-994) x POLG_MMV2T (1-455)
QY 782 AAGGAAGTGGCGGCTTCCGGGAAA 759
| | | | | | | | | | | | | | | | | |
Db 122 LysGlulLeuAlaAlaSerGlyLys 129

RESULT 12
Y456_CHLTR
ID Y456_CHLTR STANDARD; PRT; 1005 AA.
AC 084462;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein CT456 precursor.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBL_TaxID=813;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=D/UW-3/Cx;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RT Science 282:754-759(1998).
RL -1- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0572/CT456/TC0741
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC  entitled requires a license agreement (See http://www.isb-slb.ch/announce/
CC  or send an email to license@lsb-slb.ch).
CC  -----
CC  DR      EMBL; AE001319; AAC68056.1; -.
CC  DR      PHCI-2DPAGE; 084462; -.
CC  KW      Signal: Complete proteome.
CC  FT      SIGNAL      1      40
CC  FT      CHAIN       41      1005      POTENTIAL.
CC  SO      SEQUENCE   1005 AA; 102131 MW; EC47EC389851CD1E CRC64;

Alignment Scores:
Pred. No.:      8.01      Length:      1005
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.55%      Indels: 0
DB:              1      Gaps: 0

US-09-835-992A-21 (1-994) x Y456_CHLTR (1-1005)

QY      216  ACTTCAGATGCTGCTGCACACGA 239
        |||||||
DB      285  ThSerAspGlyAlaAlaAlaAla 292

RESULT 13
POLG_MNV2U      STANDARD;      PRT;      1016 AA.
ID      POLG_MNV2U
AC      P18478;
DT      01-NOV-1990 (Rel. 16, Created)
DT      01-FEB-1991 (Rel. 17, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Genome polyprotein [Contains: Nuclear inclusion protein A (NI-A) (NIA)
DE      (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear inclusion
DE      protein B (NI-B) (NIB) (RNA-directed RNA polymerase) (EC 2.7.7.48);
DE      Coat protein (CP)] (Fragment).
OS      Watermelon mosaic virus II (isolate USA).
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
OC      Polyvirus.
OX      NCBI_TaxID-148360;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Slightom J.L.;
RL      Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE OF 736-1016 FROM N.A.
RX      MEDLINE-90324931; PubMed-2374006;
RA      Quemada H., Steu L.C., Slemeniak D.R., Consalves D., Slightom J.L.;
RT      "Watermelon mosaic virus II and zucchini yellow mosaic virus: cloning
RT      of 3'-terminal regions, nucleotide sequences, and phylogenetic
RT      comparisons.";
RL      J. Gen. Virol. 71:1451-1460(1990).
CC      -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC      -1- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is
CC      further restricted by preferences for the amino acids in P6 - P1'
CC      that vary with the species of polyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
CC      Glu(Ser or Gly) for the enzyme from tobacco etch virus. The
CC      natural substrate is the viral polyprotein, but other proteins and
CC      oligopeptides containing the appropriate consensus sequence are
CC      also cleaved.
CC      -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC      (RNA)(N).
CC      -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
CC      POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC      PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC      INDIVIDUAL PROTEINS.
CC      -1- BIOTECHNOLOGY: The gene for the coat protein is introduced by
CC      genetic manipulation and expressed in squash so as to obtain virus
CC      resistant plants.
CC      -1- MISCELLANEOUS: READTHROUGH OF A TERMINATOR CODON TGA OCCURS
CC      BETWEEN RESIDUES 267-ALA AND 268-GLN.
CC      -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC      -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
CC  -----

```

```

CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; D13913; BAA03009.1; -.
DR      PIR; JQ0498; JQ0498.
DR      MEROPS; C04.003; -.
DR      InterPro; IPR001730; Peptidase_C4.
DR      InterPro; IPR001592; Poty_coat.
DR      InterPro; IPR001205; RNA_pol_P3D.
DR      Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR      Pfam; PF00767; Poty_coat; 1.
DR      Pfam; PF00863; Peptidase_C4; 1.
KW      Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KW      Coat protein; Polypeptide; Genetically modified food.
FT      NON_TER      1      1
FT      CHAIN      <1      219      NUCLEAR INCLUSION PROTEIN A.
FT      CHAIN      220      736      NUCLEAR INCLUSION PROTEIN B.
FT      CHAIN      737      1016      COAT PROTEIN.
SQ      SEQUENCE      1016 AA; 115377 MW; 8179821C4AFAFD5D CRC64;

```

```

Alignment Scores:
Pred. No.: 7.99 Length: 1016
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.58% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-21 (1-994) x POLG_MWV2U (1-1016)

QY 782 AAGGAAGTGGCGGCTTCGGGAAA 759
|||||
Db 683 LysGlutLeuAlaAlaSerGlyLys 690

RESULT 14
PAND_ECOLI
ID PAND_ECOLI STANDARD; PRT: 126 AA.
AC P31664;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Aspartate 1-decarboxylase precursor (EC 4.1.1.11) (Aspartate alpha-
DE decarboxylase).
PAND OR B0131 OR Z0142 OR ECS0135.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OC NCB1_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RA Merkel W.K., Nichols B.P.;
RL Submitted (xxx-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;

```

```

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grothbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RIMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
[5]
RN SEQUENCE OF 1-5 AND 25-29, AND CHARACTERIZATION.
RP MEDLINE=97270610; PubMed=9169598;
RX Ramjee M.K., Genschel U., Abell C., Smith A.G.;
RT "Escherichia coli L-aspartate-alpha-decarboxylase: preprotein
RT processing and observation of reaction intermediates by electrospray
RT mass spectrometry.";
RL Biochem. J. 323:661-669(1997).
[6]
RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RP MEDLINE=98206295; PubMed=9546220;
RX Albert A., Dhanaraj V., Genschel U., Khan G., Ramjee M.K., Pulido R.,
RA Sibanda B.L., von Delft F., Witty M., Blundell T.L., Smith A.G.,
RA Abell C.;
RT "Crystal structure of aspartate decarboxylase at 2.2-A resolution
RT provides evidence for an ester in protein self-processing.";
RL Nat. Struct. Biol. 5:289-293(1998).
CC -1- CATALYTIC ACTIVITY: L-aspartate = beta-alanine + CO(2).
CC -1- COFACTOR: Pyruvoyl group.
CC -1- PATHWAY: Pantothenate biosynthesis; second branch.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE PAND FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L17086; AAA24274.1; -.
DR EMBL; AE000122; AAC73242.1; -.
DR EMBL; AE005189; AAG54435.1; -.
DR EMBL; AP002550; BAB33558.1; -.
DR PDB; 1AW8; 29-APR-98.
DR EcGene; EG11747; pand.
DR InterPro; IPR003190; Asp_decarbox.
DR Pfam; PF02261; Asp_decarbox; 1.
DR TIGRFAMs; TIGR00223; pand; 1.
KW Pantothenate biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen;
KW 3D-structure; Complete proteome.
FT CHAIN 1 24 ASPARTATE 1-DECARBOXYLASE BETA CHAIN.
FT MOD_RES 25 126 ASPARTATE 1-DECARBOXYLASE ALPHA CHAIN.
FT MOD_RES 25 25 CONVERTED TO A PYRUVOYL GROUP.
SQ SEQUENCE 126 AA; 13834 MW; E3169F5C2BDD5D25 CRC64;

```

Alignment Scores:

```

Pred. No.: 127
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.23%
DB: 1

```

```

US-09-835-992A-21 (1-994) x PAND_ECOLI (1-126)
OY 647 GCAGCTGAGAGGCGAAGCCGG 667
Db 61 AlaAlaGluArgGlySerArg 67
RESULT 15
PAND_SALTY
ID PAND_SALTY STANDARD; PRT; 126 AA.
AC Q8XG14;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aspartate 1-decarboxylase precursor (EC 4.1.1.11) (Aspartate alpha-
DE decarboxylase).
GN PAND OR STM0180 OR STY0198.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson R.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
[2]
RN SEQUENCE FROM N.A.
RP SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltham D., Hamlin N., Haque A., Hien T.T., Holroyd S., Jageis K.,
RA Krogh A., Larsen T.S., Leather A., Hien T.T., Holroyd S., Jageis K.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
CC -1- CATALYTIC ACTIVITY: L-aspartate = beta-alanine + CO(2).
CC -1- COFACTOR: Pyruvoyl group (By similarity).
CC -1- PATHWAY: Pantothenate biosynthesis; second branch.
CC -1- SIMILARITY: BELONGS TO THE PAND FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008702; AAL19144.1; -.
DR EMBL; AL627265; CAD01334.1; -.
DR StyGene; SG77777; pand.
DR InterPro; IPR003190; Asp_decarbox.
DR Pfam; PF02261; Asp_decarbox; 1.
DR TIGRFAMs; TIGR00223; pand; 1.
KW Pantothenate biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen;
KW Complete proteome.
FT CHAIN 1 24 ASPARTATE 1-DECARBOXYLASE BETA CHAIN
FT CHAIN 25 126 ASPARTATE 1-DECARBOXYLASE ALPHA CHAIN

```

FT MOD_RES 25 25 (BY SIMILARITY).
FT CONVERTED TO A PYRUVOYL GROUP
SO SEQUENCE 126 AA; 13887 MW; 3A9358C00012AA80 CRC64;
(BY SIMILARITY).

Alignment Scores:

Pred. No.:	127	Length:	126
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.23%	Indels:	0
DB:	1	Gaps:	0

US-09-835-992A-21 (1-994) x PAND_SALTY (1-126)

OY 647 GCACCTGAGAGCGGAGCCGC 667
|||||
Db 61 AlaAlaGluArgGlySerArg 67

Search completed: January 14, 2003, 17:29:17
Job time : 21.349 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:20:59 : Search time 69.9914 Seconds
(without alignments)
5852.461 Million cell updates/sec

Title: US-09-835-992A-21
Perfect score: 314
Sequence: 1 ctcaccacgtgtgtctctcag.....ccaccctgggtgtgtggat 994

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+_n2p.model -DEV-xlp
-Q-/sgn2_1/USPTO/US09835992/runat_14012003_161526_4027/app_query.fasta_1.3932
-DB-SPTREMBL_21 -QFMT-fastan -SUFFIX-olig.rspt -MINMATCH-0.1 -LOOPCL-0
-LOOEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-oligo -TRANS-human40.cdl
-LIST-45 -DOCALIGN-200 -THR_SCORE-quality -THR_MIN-1 -ALIGN-15 -MODE-LOCAL
-OUTFMT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USER-US09835992_@CGN_1_1_346_@runat_14012003_161526_4027 -NCPU-6 -ICPU-3
-NO_XLIPXY -NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -LONGLOG -DEV_TIMEOUT-120
-WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60 -FGAPOP-6 -FGAPEXT-7
-YGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rv1rus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	15.9	547	11 Q9DBM7	Q9dbm7 mus musculu

2	49	15.6	405	11 Q9QW19	Q9qW19 rattus sp.
3	49	15.6	547	6 Q62742	Q62742 oryctolagus
4	30	9.6	378	5 Q9VJ44	Q9vJ44 drosophila
5	30	9.6	407	5 Q95TG8	Q95TG8 drosophila
6	30	9.6	414	5 Q24507	Q24507 drosophila
7	30	9.6	543	5 Q24506	Q24506 drosophila
8	30	9.6	544	5 Q9VJ43	Q9vJ43 drosophila
9	24	7.6	401	16 Q33332	Q33332 mycobacteri
10	22	7.0	412	5 Q02353	Q02353 caenorhabdi
11	21	6.7	400	16 Q8VJDS	Q8vJds mycobacteri
12	20	6.4	132	2 Q9EZ05	Q9eZ05 streptomyce
13	19	6.1	152	6 Q29224	Q29224 sus scrofa
14	10	3.2	387	17 Q97VJ9	Q97vJ9 sulfolobus
15	10	3.2	387	17 Q96XU9	Q96xU9 sulfolobus
16	9	2.9	246	2 Q47870	Q47870 erysipeloth
17	9	2.9	351	10 Q9ZU28	Q9zU28 arabidopsis
18	9	2.9	383	17 Q26884	Q26884 methanobact
19	9	2.9	390	17 Q97CN3	Q97cn3 thermoplasm
20	9	2.9	391	17 Q9HM39	Q9hm39 thermoplasm
21	9	2.9	510	2 Q9RA39	Q9ra39 trichodesmi
22	8	2.5	151	8 Q9MTP2	Q9mTP2 oenothera h
23	8	2.5	224	11 Q9D3L4	Q9d3L4 mus musculu
24	8	2.5	231	5 Q17964	Q17964 caenorhabdi
25	8	2.5	267	16 Q50527	Q50527 streptomyce
26	8	2.5	299	17 Q8ZTV2	Q8ztv2 pyrobaculum
27	8	2.5	308	9 Q38577	Q38577 bacterioph
28	8	2.6	314	10 Q9SZB9	Q9szb9 arabidopsis
29	8	2.5	328	5 Q9N965	Q9n965 leishmania
30	8	2.5	374	2 Q938V4	Q938v4 bradyrhizob
31	8	2.5	379	16 Q9HZQ7	Q9hzq7 pseudomonas
32	8	2.5	383	17 Q9HQZ8	Q9hqz8 halobacteri
33	8	2.5	386	16 P96838	P96838 mycobacteri
34	8	2.5	387	17 Q29811	Q29811 archaeoglob
35	8	2.5	388	17 Q29294	Q29294 archaeoglob
36	8	2.5	391	17 Q30217	Q30217 archaeoglob
37	8	2.5	394	16 Q99WZ2	Q99wz2 staphylococ
38	8	2.5	394	17 Q8TVL1	Q8tvL1 methanopyru
39	8	2.5	395	17 Q97VV9	Q97vv9 sulfolobus
40	8	2.5	395	17 Q8ZV25	Q8zv25 pyrobaculum
41	8	2.5	396	17 Q978T8	Q978t8 thermoplasm
42	8	2.5	396	17 Q9HLD2	Q9hld2 thermoplasm
43	8	2.5	397	2 Q93SM1	Q93sm1 staphylococ
44	8	2.5	399	16 Q8Y261	Q8y261 raistonia s
45	8	2.5	400	16 Q9A8C2	Q9a8C2 caulobacter

ALIGNMENTS

RESULT 1	ID	Q9DBM7	PRELIMINARY;	PRT;	547 AA.
Q9DBM7	Q9DBM7	01-JUN-2001 (TREMBLrel. 17, Created)			
AC	Q9DBM7	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)			
DE	Sterol carrier protein 2, liver.				
GN	SCP2.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=LIVER;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,				
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Queckenbush J.,				
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
RN [2]

RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK004860; BAB23623.1; -
DR EMBL; BC018384; AAH18384.1; -
DR MGD; MGI:98254; SCP2.
DR InterPro; IPR003033; SCP2.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF02036; SCP2; 1.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLEASE_1; 1.
DR PROSITE; PS00737; THIOLEASE_2; 1.
SQ SEQUENCE 547 AA; 59125 MW; 37BA2E730D9CB105 CRC64;

Alignment Scores:
Pred. No.: 6.52e-45 Length: 547
Score: 50.00 Matches: 50
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.92% Indels: 0
DB: 11 Gaps: 0

US-09-835-992A-21 (1-994) x Q9QW19 (1-547)

QY 522 ACGCTGTTGATAGAGAGATAATACATATGAGAGAAAGTGGTCAATAATCCTAGTGT 581
Db 329 ThrleuValaspargGlyaspasnThrTyrglyGlyLysTrpValIleasnProserGly 348
QY 582 GGAAGTTCACAAAGGACACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGACAGACTC 641
Db 349 GlyleuIleSerLysGlyHisProleuGlyAlaThrGlyLeuAlaGlnCysAlaGluLeu 368
QY 642 TGGTGGCAGCTGAGAGGGGGAAGCCGGA 671
Db 369 CysTrpGlnLeuArgGlyGluAlaGlyLys 378

RESULT 2

Q9QW19 PRELIMINARY; PRT; 405 AA.
AC Q9QW19;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE 3-oxoacyl-CoA thiolase homolog (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10118;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93283441; PubMed=8507681;
RA Ossendorp B.C., Wirtz K.W.;
RT "The non-specific lipid-transfer protein (sterol carrier protein 2)
RT and its relationship to peroxisomes";
RL Biochimie 75:191-200(1993).
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLEASE_1; 1.
DR PROSITE; PS00737; THIOLEASE_2; 1.

DR PROSITE; PS00737; THIOLEASE_2; 1.
SQ SEQUENCE 405 AA; 43650 MW; 8C6782095D7F628F CRC64;

Alignment Scores:
Pred. No.: 8.26e-44 Length: 405
Score: 49.00 Matches: 49
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.61% Indels: 0
DB: 11 Gaps: 0

US-09-835-992A-21 (1-994) x Q9QW19 (1-405)

QY 525 CTGGTTGATAGAGAGATATATGAGAGAAAGTGGTCAATAATCCTAGTGTGGA 584
Db 329 LeuValaspargGlyaspasnThrTyrglyGlyLysTrpValIleasnProserGly 348
QY 585 CTGATTTCAAAGGACACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGACAGACTTCGC 644
Db 349 LeuIleSerLysGlyHisProleuGlyAlaThrGlyLeuAlaGlnCysAlaGluLeu 368
QY 645 TGGCAGCTGAGAGGGGAAGCCGGA 671
Db 369 TrpGlnLeuArgGlyGluAlaGlyLys 377

RESULT 3

O62742 PRELIMINARY; PRT; 547 AA.
ID O62742;
AC O62742;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Sterol carrier protein X.
GN SCP2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Weber F.E., Werder M., Boffelli D., Hauser H.;
RT "In pre-sterol carrier protein 2 (SCP2) in solution the leader peptide
RT 1-20 is flexibly disordered and the residues 21-143 adopt the same
RT globular fold as in mature SCP2";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF051897; AAC15422.1; -
DR HSSP; P27796; 1PXT.
DR InterPro; IPR003033; SCP2.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF02036; SCP2; 1.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLEASE_1; 1.
DR PROSITE; PS00737; THIOLEASE_2; 1.
SQ SEQUENCE 547 AA; 58903 MW; 32BE7F7DCBF9BCD7 CRC64;

Alignment Scores:
Pred. No.: 7.97e-44 Length: 547
Score: 49.00 Matches: 49
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.61% Indels: 0
DB: 6 Gaps: 0

US-09-835-992A-21 (1-994) x O62742 (1-547)

QY 525 CTGGTTGATAGAGAGATATATGAGAGAAAGTGGTCAATAATCCTAGTGTGGA 584
Db 330 LeuValaspargGlyaspasnThrTyrglyGlyLysTrpValIleasnProserGly 349
QY 585 CTGATTTCAAAGGACACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGACAGACTTCGC 644
Db 350 LeuIleSerLysGlyHisProleuGlyAlaThrGlyLeuAlaGlnCysAlaGluLeu 369

OY 645 TGGCAGCTGACAGGCGAAGCGGAAA 671
Db 370 TrpGlnLeuArgGlyGlnAglGlyLys 378

RESULT 4

09VJ44 PRELIMINARY; PRT; 378 AA.
AC 09VJ44;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE CG17597 protein.
GN CG17597.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

SEQUENCE FROM N.A.
STRAIN=BERKELEY;
MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sulton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlner C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hosteln D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Llu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003660; AAF53712.1; -
DR FLYBase; FBgn0032715; CG17597.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLEASE_1; 1.
DR PROSITE; PS00737; THIOLEASE_2; 1.
SQ SEQUENCE 378 AA; 40655 MW; 704CB62E7E6F075E CRC64;

Alignment Scores: 3.76e-23 Length: 378
Pred. No.: 30.00 Matches: 30
Score:

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.55% Indels: 0
DB: 5 Gaps: 0

US-09-835-992A-21 (1-994) x Q9VJ44 (1-378)

QY 570 AATCCTAGTGGTGACTGATTTCAGAGGACACCCACTAGCGCCTACAGGCTCTGCTCAG 629
Db 308 AsnProserGlyGlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 327
QY 630 TGTGCAACTCTGCTGGCAGCTGACAGCG 659
Db 328 CysAlaGluLeuCysTrpGlnLeuArgGly 337

RESULT 5

Q95TG8 PRELIMINARY; PRT; 407 AA.
AC Q95TG8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE SD06839P.
GN CG17597.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY058779; AAL14008.1; -
DR FLYBase; FBgn0032715; CG17597.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLEASE_1; UNKNOWN_1.
DR PROSITE; PS00737; THIOLEASE_2; UNKNOWN_1.
SQ SEQUENCE 407 AA; 43904 MW; 7BE64B5D68C49D1B CRC64;

Alignment Scores:

Pred. No.: 3.73e-23 Length: 407
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.55% Indels: 0
DB: 5 Gaps: 0

US-09-835-992A-21 (1-994) x Q95TG8 (1-407)

QY 570 AATCCTAGTGGTGACTGATTTCAGAGGACACCCACTAGCGCCTACAGGCTCTGCTCAG 629
Db 337 AsnProserGlyGlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 356
QY 630 TGTGCAACTCTGCTGGCAGCTGACAGCG 659
Db 357 CysAlaGluLeuCysTrpGlnLeuArgGly 366

RESULT 6

Q24507 PRELIMINARY; PRT; 414 AA.
AC Q24507;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Sterol carrier protein X-related thiolase.
GN SCPX OR CG17320.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Kitamura T., Kobayashi S., Okada M.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X95443; CAA64718.1; -.
DR FlyBase; FBgn0015808; Scpx.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLASE_1; 1.
DR PROSITE; PS00737; THIOLASE_2; 1.
SQ SEQUENCE 414 AA; 44750 MW; 40CFAFE836111B7E CRC64;

Alignment Scores:
Pred. No.: 3.72e-23 Length: 414
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.55% Indels: 0
DB: 5 Gaps: 0

US-09-835-992A-21 (1-994) x Q24507 (1-414)

QY 570 AATCCTAGTGGTGAAGGACCCACTAGCGGCTACAGGCTCTGCTCAG 629
|||||
Db 337 AsnProSerGlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 356

QY 630 TGTGACGAACCTGCTGGCAGCTGAGAGG 659
|||||
Db 357 CysAlaGluLeuGlyStrpGlnLeuArgGly 366

RESULT 7

Q24506 PRELIMINARY; PRT; 543 AA.
ID Q24506;
AC Q24506;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE Sterol carrier protein X.
GN SCPX OR CG17320.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Kitamura T., Kobayashi S., Okada M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X97685; CAA66277.1; -.
DR FlyBase; FBgn0015808; Scpx.
DR InterPro; IPR003033; SCP2.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF02036; SCP2; 1.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLASE_1; 1.
DR PROSITE; PS00737; THIOLASE_2; 1.
SQ SEQUENCE 543 AA; 58854 MW; 4B9BB43E77048181 CRC64;

Alignment Scores:
Pred. No.: 3.61e-23 Length: 543
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.55% Indels: 0
DB: 5 Gaps: 0

us-09-835-992A-21 (1-994) x Q24506 (1-543)

QY 570 AATCCTAGTGGTGAAGGACCCACTAGCGGCTACAGGCTCTGCTCAG 629
|||||
Db 337 AsnProSerGlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 356

RESULT 8

Q9VJ43 PRELIMINARY; PRT; 544 AA.
ID Q9VJ43;
AC Q9VJ43;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE CG17320 protein (LD10783p).
GN SCPX OR CG17320.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Rulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Splier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weltsenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).

Q9VJ43 SEQUENCE FROM N.A.
RP STRAIN=BERKELEY;
RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleob J., Paragas V., Park S.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Colinker S.,
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003660; AAF53713.1; -.
DR EMBL; AY089539; AAL90277.1; -.
DR FlyBase; FBgn0015808; ScpX.
DR InterPro; IPR003033; SCP2.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF02036; SCP2; 1.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLASE_1; 1.
DR PROSITE; PS00737; THIOLASE_2; 1.
SQ SEQUENCE 544 AA; 59007 MW; A14B0A7D2185E9C2 CRC64;

Alignment Scores:
Pred. No.: 3.61e-23 Length: 544
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.55% Indels: 0
DB: 5 Gaps: 0

US-09-835-992A-21 (1-994) x Q9VJ43 (1-544)

QY 570 AATCCTAGTGGTGGACTGATTTCAAAGGACACCCACTAGCGCTACAGTCTTGTCTCAG 629
Db 337 AsnProSerGlyGlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 356
QY 630 TGTGCAGAACTCTGCTGCGCAGCTGACAGCG 659
Db 357 CysAlaGluLeuCysTrpGlnLeuArgGly 366

RESULT 9
ID 033332 PRELIMINARY; PRT; 401 AA.

AC 033332;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Lipid-transfer protein.
GN LTP1 OR RV2790C OR MTV002.55C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham D., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL; AL008967; CAA15585.1; -.
DR HSSP; P27796; IAFY.
DR Tuberculatc; RV2790C; -.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLASE_1; 1.
DR PROSITE; PS00737; THIOLASE_2; 1.
KW Complete proteome.
SQ SEQUENCE 401 AA; 42904 MW; 6059E0BB995C607E CRC64;

Alignment Scores:

Pred. No.: 1.25e-16 Length: 401
Score: 24.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.64% Indels: 0
DB: 16 Gaps: 0

US-09-835-992A-21 (1-994) x O02353 (1-412)

QY 570 AATCCTAGTGGTGGACTGATTTCAAAGGACACCCACTAGCGCTACAGTCTTGTCTCAG 629
Db 338 AsnProSerGlyGlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 357
QY 630 TGTGCAGAACTC 641
Db 358 CysAlaGluLeu 361

RESULT 10

ID 002353 PRELIMINARY; PRT; 412 AA.

AC 002353;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE 3-keto-acyl-CoA thiolase.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-N2;
RX MEDLINE=97296386; PubMed=9151950;
RA Bun-Ya M., Maebuchi M., Hashimoto T., Yokota S., Kamiryo T.;
RT "A second isoform of 3-ketoacyl-CoA thiolase found in Caenorhabditis
RT elegans, which is similar to sterol carrier protein 2.";
RT sequence of sterol carrier protein 2.";
RL Eur. J. Biochem. 245:252-259(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Smye R.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; D86473; BAA20377.1; -.
DR EMBL; AL023847; CAA19548.1; -.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLASE_1; UNKNOWN_1.
DR PROSITE; PS00737; THIOLASE_2; 1.
SQ SEQUENCE 412 AA; 44386 MW; 6B9DA9BA18ABF446 CRC64;

Alignment Scores:
Pred. No.: 1.85e-14 Length: 412
Score: 22.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.01% Indels: 0
DB: 5 Gaps: 0

US-09-835-992A-21 (1-994) x O02353 (1-412)

QY 540 GATAATACATATGAGGAAGTGGCTTAATAATCTAGTGGTGGACTGATTTCAAAGGCA 599
Db 332 AspAsnThrTyGlyGlyLysTrpValIleAsnProSerGlyGlyLeuIleSerLysGly 351
QY 600 CACCCA 605
Db 352 HisPro 353

RESULT 11

ID 08VJD5 PRELIMINARY; PRT; 400 AA.
AC 08VJD5;

DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Nonspecific lipid-transfer protein.
GN MT2860.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE007112; AAK47179.1; -.
DR TIGR; MT2860; -.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00108; thiolase; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLEASE_1; UNKNOWN_1.
DR PROSITE; PS00737; THIOLEASE_2; UNKNOWN_1.
SQ SEQUENCE 400 AA; 42833 MW; 8A75C1BB204DF7A2 CRC64;

Alignment Scores:

Pred. No.:	2.27e-13	Length:	400
Score:	21.00	Matches:	21
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.69%	Indels:	0
DB:	16	Gaps:	0

US-09-835-992A-21 (1-994) x Q8VJD5 (1-400)

QY 570 AATCCTAGTGTGAGCTGATTTCAAAGGAGACCCACTAGGCGCTACAGTCTTGCTCAG 629
|||||
DB 338 AsnProserGlyGlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 357

QY 630 TGT 632
|||

DB 358 Cys 358

RESULT 12

Q9EZ05 PRELIMINARY; PRT; 132 AA.

AC Q9EZ05;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 13.3 kDa protein (Fragment).
OS Streptomyces aureofaciens.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1894;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM3239;
RA Kormanec J., Novakova R., Homerova D., Rezuchova B.;
RT "Putative meta-cleavage compound hydrolase gene is dependent on a
RT sporulation specific sigma RpoZ in Streptomyces aureofaciens.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF306830; AAG42771.1; -.
DR HSSP; P27796; 1AFY.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00737; THIOLEASE_2; 1.
KW Hypothetical protein.
FT NON_TER 1

SQ SEQUENCE 132 AA; 13296 MW; E1173B31584DE487 CRC64;
Alignment Scores:
Pred. No.: 3.17e-12 Length: 132
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.37% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x Q9EZ05 (1-132)

QY 570 AATCCTAGTGTGAGCTGATTTCAAAGGAGACCCACTAGGCGCTACAGTCTTGCTCAG 629
|||||
DB 72 AsnProserGlyGlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGln 91

RESULT 13

Q29224 PRELIMINARY; PRT; 152 AA.

AC Q29224;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Sterol carrier protein-X/sterol carrier protein-2 (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RX MEDLINE=96327607; PubMed=8672129;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
RL library.";
DR EMBL; F14816; CAA23274.1; -.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF02803; thiolase_C; 1.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 152
SQ SEQUENCE 152 AA; 16276 MW; 9632305FE9FC8A6A CRC64;

Alignment Scores:

Pred. No.:	3.81e-11	Length:	152
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.05%	Indels:	0
DB:	6	Gaps:	0

US-09-835-992A-21 (1-994) x Q29224 (1-152)

QY 363 GTTGCTTTGATATGAGTAAGAGCTGCAGAAATGCTATGAGAAATCTGGCCTG 419
|||||
DB 72 ValGlyPheAspMetSerLysGluAlaAlaArgLysCysTyrGluLysSerGlyLeu 90

RESULT 14

Q97VJ9 PRELIMINARY; PRT; 387 AA.

AC Q97VJ9;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Acetyl-CoA c-acetyltransferase (acetoacetyl-CoA thiolase)
GN ACAB-7 OR SSO2625.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE-21332296; PubMed-11427726;
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweye M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moers A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL: AE006858; AAK42745.1; -.
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF00108; thiolase; 1.
DR Pfam: PF02803; thiolase_C; 1.
KW Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 387 AA; 41892 MW; 4F6E505FEDB3522F CRC64;

Alignment Scores:
Pred. No.: 0.207 Length: 387
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.18% Indels: 0
DB: 17 Gaps: 0

US-09-835-992A-21 (1-994) x Q97VJ9 (1-387)

OY 594 AAGGACACCCACTAGCGCTACAGTCTT 623
Db 333 LysGlyHisProLeuGlyAlaThrGlyLeu 342

RESULT 15

ID Q96XU9 PRELIMINARY; PRT; 387 AA.
AC Q96XU9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative nonspecific lipid-transfer protein.
GN ST2418.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID-111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JCM 10545 / 7;
RX PubMed-11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL: AP000990; BAB67528.1; -.
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF00108; thiolase; 1.
DR Pfam: PF02803; thiolase_C; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 387 AA; 42071 MW; DF3372F987C42D69 CRC64;

Alignment Scores:
Pred. No.: 0.207 Length: 387
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.18% Indels: 0
DB: 17 Gaps: 0

US-09-835-992A-21 (1-994) x Q96XU9 (1-387)

OY 594 AAGGACACCCACTAGCGCTACAGTCTT 623
Db 333 LysGlyHisProLeuGlyAlaThrGlyLeu 342

Search completed: January 14, 2003, 17:37:03
Job time : 77.9914 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:24 ; Search time 41.5252 Seconds

(without alignments)
5410.226 Million cell updates/sec

Title: US-09-835-992A-22

Perfect score: 180

Sequence: 1 ggcccaaaaantattttna.....taatccgacaataaaaag 843

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+_n2p.model -DEV-rlp
-Q-/cgn2.1/USPTO_spool/US09835992/runat_14012003_161525_3979/app_query.fasta_1.3932
-DB-A_Geneseq_101002 -QFMT-fastan -SUFFIX-olig.rag -MINMATCH-0.1 -LOOPCL-0
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-oligo -TRANS-human40.cdl
-LIST-45 -DOCALIGN-200 -THR_SCORE-quality -THR_MIN-1 -ALIGN-15 -MODE-LOCAL
-OUTFMT=ptc -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USER=US09835992@cgn_1_1_121_erunat_14012003_161525_3979 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -LONGLOG -DEV_TIMEOUT-120
-WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60 -FGAPOP-6 -FGAPEXT-7
-YGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database :

A_Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
C	1	7	3.8	56	20	AAV02476	Clone selected aft
	2	7	3.9	59	23	ABP05819	Human ORFX protein
	3	7	3.9	172	22	AAB52473	Mycobacterium tube
C	4	7	3.8	185	21	AAV74955	Neisseria meningit
	5	7	3.9	188	20	AAW95491	M. tuberculosis RP
	6	7	3.9	584	22	ABG05212	Novel human diagno
	7	7	3.9	594	22	ABG21682	Novel human diagno
	8	7	3.9	691	21	AAB07698	A Neisseria mening
	9	7	3.9	722	21	AAB07697	A Neisseria mening
	10	7	3.9	1718	22	ABG24524	Novel human diagno
	11	6	3.3	15	21	AAB03322	Human epitope Rh50
	12	6	3.3	25	21	AAV69422	Antimicrobial pept
	13	6	3.3	25	21	AAV69423	Antimicrobial pept
	14	6	3.3	25	21	AAV69424	Antimicrobial pept
	15	6	3.3	25	21	AAV69469	Antimicrobial pept
	16	6	3.3	35	16	AAR74455	Parathyroid hormon
	17	6	3.3	35	22	ABG01026	Novel human diagno
C	18	6	3.3	36	23	ABP30830	Streptococcus poly
C	19	6	3.3	39	21	AAB38159	Human secreted pro
	20	6	3.3	41	21	AAB44645	Human secreted pro
	21	6	3.3	42	21	AAB12320	Human gene 1 encod
	22	6	3.3	43	23	AAE21238	Human immune/haema
C	23	6	3.3	46	22	AAM90896	Human peptide enco
C	24	6	3.3	46	23	ABG46482	Human ORFX protein
	25	6	3.3	50	23	ABP04160	Human ORFX protein
C	26	6	3.3	57	22	ABG12669	Novel human diagno
	27	6	3.3	58	22	AAO04926	Human polypeptide
	28	6	3.3	61	22	ABB14799	Human nervous syst
	29	6	3.3	64	22	AAO03024	Human polypeptide
C	30	6	3.3	65	22	AAU51896	Propionibacterium
C	31	6	3.3	65	23	AAU80965	Toxoplasma Graz an
C	32	6	3.3	67	23	ABP01643	Human ORFX protein
C	33	6	3.3	68	23	ABP04033	Human ORFX protein
	34	6	3.3	72	23	ABG60075	Human DITHP polype
	35	6	3.3	73	21	AAB40726	Human ORFX ORF490
C	36	6	3.3	74	22	AAU86612	Novel human connec
C	37	6	3.3	74	22	AAO00202	Human polypeptide
C	38	6	3.3	78	22	AAO03840	Human polypeptide
	39	6	3.3	78	23	ABP03178	Human ORFX protein
	40	6	3.3	79	23	ABP42461	Human ovarian anti
C	41	6	3.3	85	23	ABP33738	Human ORF2711 prot
	42	6	3.3	86	20	AAV07953	Human secreted pro
C	43	6	3.3	86	20	AAV07852	Human secreted pro
C	44	6	3.3	86	22	AAM83146	Human immune/haema
C	45	6	3.3	87	22	AAO03194	Human polypeptide

ALIGNMENTS

RESULT 1
AAY02476
ID AAY02476 standard; Protein; 56 AA.
XX
AC AAY02476;
XX
DT 15-JUL-1999 (first entry)
XX
DE Clone selected after panning a NK library of the invention.
XX Screening; functional polypeptide; ligand; non-functional;
KM enrichment; single chain antibody; Scfv.
XX
OS Synthetic.
XX
XX WO9920749-A1.
XX PN
XX 29-APR-1999.
XX

PF 20-OCT-1998; 98WO-GB03135.
XX
PR 21-NOV-1997; 97US-0066729.
PR 20-OCT-1997; 97GB-0022131.
PR 13-NOV-1997; 97US-0065428.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Tomlinson I, Winter G;
XX
DR WPI; 1999-288302/24.
XX
PT Screening for functional polypeptides which bind a ligand
XX
PS Example 3; Fig 4; 67pp; English.
XX
CC The specification describes a method for screening for functional
CC polypeptides which bind a ligand. The method comprises contacting a
CC repertoire of polypeptides with a generic ligand, and then screening a
CC selected functional polypeptides with a target ligand. The method
CC permits the removal from a chosen repertoire of polypeptides, those
CC which are non-functional, e.g. as a result of the introduction of
CC frame-shift mutations, stop codons, folding mutants or expression
CC mutants which would be or are incapable of binding to any target
CC ligand. The method also permits the enrichment of a chosen repertoire
CC of polypeptides for those polypeptides which are functional, well folded
CC and highly expressed. The polypeptides obtained can be used in
CC diagnostic, prophylactic and therapeutic procedures. AAY02473-Y02525
CC represent clones selected after panning primary and somatic NMK
CC libraries with 5 antigens (bovine ubiquitin, rat BIP, bovine histone,
CC NIP-BSA and hen egg lysozyme).
XX
SQ Sequence 56 AA;

Alignment Scores:
Pred. No.: 49.5 Length: 56
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.83% Indels: 0
DB: 20 Gaps: 0

US-09-835-992A-22 (1-843) x AAY02476 (1-56)
QY 802 AAAGNGGGGAGGTGTTT 782
Db 21 LysGlyGlyGlyArgLeuPhe 27

RESULT 2
ABP05819
ID ABP05819 standard; Protein; 59 AA.
XX
AC ABP05819;
XX
DT 24-JUN-2002 (first entry)
XX
DE Human ORFX protein sequence SEQ ID NO:11620.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX

PF 29-MAY-2001; 2001WO-US10836.
XX
PR 30-MAY-2000; 2000US-206132P.
PR 29-AUG-2000; 2000US-228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach MD;
XX
DR WPI; 2002-106308/14.
DR N-PSDB; ABN21571.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders
XX
PS Disclosure; SEQ ID 11620; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 59 AA;

Alignment Scores:
Pred. No.: 49.1 Length: 59
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 23 Gaps: 0

US-09-835-992A-22 (1-843) x ABP05819 (1-59)
QY 773 CCCCTTAAAAACAACCTTCC 793
Db 34 ProLeuLysLysGlnProSer 40

RESULT 3
AAB52473
ID AAB52473 standard; protein; 172 AA.
XX
AC AAB52473;
XX
DT 23-FEB-2001 (first entry)
XX
DE Mycobacterium tuberculosis secreted protein #38.
XX
KW Mycobacterium tuberculosis secreted protein; MTSP; vaccine.
XX
OS Mycobacterium tuberculosis.

XX PN WO200066143-A1.
XX PD 09-NOV-2000.
XX PF 04-MAY-2000; 2000WO-US12197.
XX PR 04-MAY-1999; 99US-0132479.
XX PR 04-MAY-1999; 99US-0132503.
XX PA (PUBL-) PUBLIC HEALTH RES INST NEW YORK.
XX PI Gennaro ML, Gomez MJ;
XX DR WPI; 2001-007151/01.
XX PT Novel Mycobacterium tuberculosis secreted polypeptides and
PT polynucleotides useful in diagnosis, treatment and prophylaxis of
PT tuberculosis -
XX PS Claim 11; Fig 1; 60pp; English.
XX CC The present invention relates to Mycobacterium tuberculosis secreted
CC proteins (MTSP), where the polypeptide has M. tuberculosis specific
CC antigenic and immunogenic properties. Compositions of the invention may
CC be useful for diagnosing Mycobacterium tuberculosis infection and as a
CC vaccine against M. tuberculosis infection.
XX SQ Sequence 172 AA;

Alignment Scores:
Pred. No.: 41.7 Length: 172
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-22 (1-843) x AAB52473 (1-172)
QY 786 AACCTTCCCCNCCTTTGGCC 806
Db 73 AsnLeuProProLeuAla 79

RESULT 4
AAV74955
ID AAV74955 standard; Protein; 185 AA.
XX AC AAV74955;
XX DT 21-MAR-2000 (first entry)
XX DE Neisseria meningitidis ORF 507 protein sequence SEQ ID NO:1384.
XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX OS Neisseria meningitidis.
XX PN WO9957280-A2.
XX PD 11-NOV-1999.
XX PF 30-APR-1999; 99WO-US09346.
XX PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.

PR 25-FEB-1999; 99US-0121528.
XX PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
PI Petersen J, Pizze M, Rappoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX DR WPI; 2000-062150/05.
DR N-PSDB; AAZ53717.
XX PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX PS Claim 2; Page 752-753; 1453pp; English.
XX CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAV74253 to AAV75941
CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX SQ Sequence 185 AA;

Alignment Scores:
Pred. No.: 41.2 Length: 185
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.83% Indels: 0
DB: 21 Gaps: 0

US-09-835-992A-22 (1-843) x AAV74955 (1-185)
QY 789 GGTGTTTTTTAAGGGGGA 769
Db 10 GlyCysPheLeuArgGlyGly 16

RESULT 5
AAW95491
ID AAW95491 standard; Protein; 188 AA.
XX AC AAW95491;
XX DT 26-MAR-1999 (first entry)
XX DE M. tuberculosis RP-factor-like gene product (MtubMTV008).
XX KW RP-factor; resuscitate; latent cell; growth-stimulation; receptor;
KW convertase; vaccine; immunotherapy; prophylaxis; excipient; antibiotic;
KW antimicrobial; tuberculosis; food; pharmaceutical; culture additive.
XX OS Mycobacterium tuberculosis.
XX PN WO9855624-A1.
XX PD 10-DEC-1998.
XX PF 03-JUN-1998; 98WO-GB01619.
XX PR 27-MAY-1998; 98GB-0011221.
PR 04-JUN-1997; 97GB-0011389.
XX PA (UYWA-) UNIV WALES.

XX	Kaprelyants AS,	Kell DB,	Mukamlova GV,	Young DI;
PI	Young M;			
XX				
DR	WPI; 1999-070270/06.			
XX				
PT	New bacterial resuscitation factors - useful for stimulating latent			
PT	bacteria and growth, particularly for use as antimicrobials, also			
PT	vaccines comprising bacteria with mutations in resuscitation factor			
PT	genes			
XX				
PS	Disclosure; Fig 1A; 76pp; English.			
XX				
CC	The invention relates to RP-factors which are substances that can			
CC	resuscitate dormant, moribund or latent cells, possibly also having			
CC	growth-stimulating activity. Host cells containing a vector comprising			
CC	the RP-factor or its receptor encoding nucleic acid can be used for the			
CC	recombinant production of the RP-factor. RP-factors, their receptors or			
CC	convertases, antibodies (against the RP-factors and RP-factor receptors			
CC	or convertases), antagonists and agonists, are useful in vaccines and for			
CC	immunotherapy, diagnosis and prophylaxis. They are also useful as			
CC	excipients, generally as antimicrobials and especially for infections			
CC	associated with latency. They can be used as potentiators of antibiotics			
CC	such as isoniazid, streptomycin etc., in treatment of tuberculosis.			
CC	RP-factor can also be used: to determine the microbiological quality of			
CC	foods, pharmaceuticals, medical products; as culture additive for			
CC	bacteria; to stimulate growth and/or to resuscitate microorganisms and to			
CC	produce libraries of biomolecules and microorganisms (which may then be			
CC	screened for useful products). Probes complementary to the RP-factor			
CC	nucleic acid are used identify and clone other RP-factor genes. The			
CC	inventions may also be used to prevent bacterial resuscitation. Breaking			
CC	dormancy with RP-factor facilitates detection, culture and enumeration of			
CC	many bacteria. The present sequence represents a predicted amino acid			
CC	sequence of a RP-factor-like gene product from M. tuberculosis.			
XX				
SQ	Sequence	188 AA;		
Alignment Scores:				
Pred. No.:	41.1	Length:	188	
Score:	7.00	Matches:	7	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	3.89%	Indels:	0	
DB:	20	Gaps:	0	
US-09-835-992A-22 (1-843) x AAW95491 (1-188)				
OY	786 AACCTTCCCCCNCCTTTGGCC	806		
Db	89 AsnleuProProProlenuAa	95		
RESULT 6				
ABG05212	standard; Protein; 584 AA.			
ID	ABG05212			
XX	AC ABG05212;			
XX				
DT	13-FEB-2002	(first entry)		
XX				
DE	Novel human diagnostic protein #5203.			
XX				
KM	Human; chromosome mapping; gene mapping; gene therapy; forensic;			
KW	food supplement; medical imaging; diagnostic; genetic disorder.			
XX				
OS	Homo sapiens.			
XX				
PN	WO200175067-A2.			
XX				
PD	11-OCT-2001.			
XX				
PF	30-MAR-2001; 2001WO-US08631.			
XX				
PR	31-MAR-2000; 2000US-0540217.			

```

PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX N-PSDB; AAS69399.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity
XX
XX Claim 20; SEQ ID No 35571; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 584 AA;
XX
XX
XX Alignment Scores:
XX Pred. No.: 34.6 Length: 584
XX Score: 7.00 Matches: 7
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 3.89% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-835-992A-22 (1-843) x ABG05212 (1-584)
XX
XX QY 670 TTACCA TTTTACCTGTGTAAG 690
XX |||||||
XX Db 172 LeuProPheTyrLeuGlyLys 178
XX
XX RESULT 7
XX ABG21682
XX ID ABG21682 standard; Protein; 594 AA.
XX
XX AC ABG21682;
XX
XX DT 18-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #21673.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX

```

PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85869.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID NO 52041; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 594 AA;

Alignment Scores:
Pred. No.: 34.5 Length: 594
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-22 (1-843) x ABG21682 (1-594)

QY 670 TTACCATTTTACCTTGCTAAG 690
XX ||||||||||||||||
DB 172 LeuProPheTyrLeuGlyLys 178

RESULT 8
AAB07698
ID AAB07698 standard; Protein; 691 AA.
XX
AC AAB07698;
XX
DT 07-NOV-2000 (first entry)
XX
DE A Neisseria meningitidis BASB053 polypeptide.
XX
KW BASB053; Neisseria meningitidis infection; vaccine.
XX
OS Neisseria meningitidis.
XX
PN WO200042193-A1.
XX

PD 20-JUL-2000.
XX
XX
PF 10-JAN-2000; 2000WO-EP00137.
XX
PR 15-JAN-1999; 99GB-0000959.
PR 28-JAN-1999; 99GB-0001903.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI; 2000-476062/41.
DR N-PSDB; AAA59217.
XX
PT New Neisseria meningitidis polypeptide useful for diagnosis of
PT Neisseria infection and for development of vaccines against such
PT infection -
XX
XX
PS Claim 3; Page 56; 92pp; English.
XX
CC The present sequence represents a Neisseria meningitidis BASB053
CC polypeptide. The BASB053 polypeptide, or an antibody immunospecific
CC for BASB053 may be identified in a biological sample in order to
CC diagnose a Neisseria meningitidis infection in an animal. The BASB053
CC polypeptides and polynucleotides may be used as vaccines, for
CC generating an immune response in an animal. A composition comprising
CC at least one antibody immunospecific for BASB053 may be used to
CC treat humans infected with Neisseria meningitidis.
XX
SQ Sequence 691 AA;

Alignment Scores:
Pred. No.: 33.7 Length: 691
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 21 Gaps: 0

US-09-835-992A-22 (1-843) x AAB07698 (1-691)

QY 823 CTAATCCGAACAATAAAG 843
XX ||||||||||||||||
DB 535 LeuAsnProAsnAsnLysLys 541

RESULT 9
AAB07697
ID AAB07697 standard; Protein; 722 AA.
XX
AC AAB07697;
XX
DT 07-NOV-2000 (first entry)
XX
DE A Neisseria meningitidis BASB053 polypeptide.
XX
KW BASB053; Neisseria meningitidis infection; vaccine.
XX
OS Neisseria meningitidis.
XX
PN WO200042193-A1.
XX
PD 20-JUL-2000.
XX
PF 10-JAN-2000; 2000WO-EP00137.
XX
PR 15-JAN-1999; 99GB-0000959.
PR 28-JAN-1999; 99GB-0001903.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI; 2000-476062/41.
XX

DR N-PSDB; AAA59216.
XX
PT New Neisseria meningitidis polypeptide useful for diagnosis of
PT Neisseria infection and for development of vaccines against such
PT infection -
XX
PS Claim 3; Page 55-56; 92pp; English.
XX
CC The present sequence represents a Neisseria meningitidis BASB053
CC polypeptide. The BASB053 polypeptide, or an antibody immunospecific
CC for BASB053 may be identified in a biological sample in order to
CC diagnose a Neisseria meningitidis infection in an animal. The BASB053
CC polypeptides and polynucleotides may be used as vaccines, for
CC generating an immune response in an animal. A composition comprising
CC at least one antibody immunospecific for BASB053 may be used to
CC treat humans infected with Neisseria meningitidis.
XX
SQ Sequence 722 AA;

Alignment Scores:
Pred. No.: 33.5 Length: 722
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 21 Gaps: 0

US-09-835-992A-22 (1-843) x AAB07697 (1-722)

OY 823 CTAATCCGACATATAAAG 843
Db 566 LeuAsnProAsnAsnLys 572

RESULT 10
ABG24524
ID ABG24524 standard; Protein; 1718 AA.
XX
AC ABG24524;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #24515.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS88711.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 54883; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1718 AA;

Alignment Scores:
Pred. No.: 29.4 Length: 1718
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 22 Gaps: 0

US-09-835-992A-22 (1-843) x ABG24524 (1-1718)

OY 670 TTACCATTTTACCTTGTAAG 690
Db 220 LeuProPheTyrLeuGlyLys 226

RESULT 11
AAB03322
ID AAB03322 standard; Peptide; 15 AA.
XX
AC AAB03322;
XX
DT 08-SEP-2000 (first entry)
XX
DE Human epitope Rh50 GP # 32 corresponding to residues 311-325.
XX
KW Human; rhesus blood group system; Rh; RhD; RhCE; sickle cell disease;
KW thalassaemia; RhC; Rhc; RHe; Rhe; alloimmunisation prevention;
KW autoimmune Rh haemolytic disease; rhesus protein; immunosuppressive;
KW vaccine.
XX
OS Homo sapiens.
XX
PN WO200032632-A2.
XX
PD 08-JUN-2000.
XX
PF 01-DEC-1999; 99WO-GB04027.
XX
PR 01-DEC-1998; 98GB-0026378.
XX
PA (UYAB-) UNIV ABERDEEN.
PA (COMM-) COMMON SERVICES AGENCY SCOTTISH HEALTH S.
XX
PI Urbaniaak SU, Barker RN;
XX
DR WPI; 2000-412291/35.
XX
PT Composition for prevention of alloimmunization or immunosuppression of
PT a response elicited by alloimmunization or an autoimmune hemolytic
PT disease, comprises an epitope of a rhesus protein -
XX
PS Disclosure; Page 78; 92pp; English.

XX Human blood contains the rhesus (Rh) blood group system, and humans can
CC either be RhD positive or negative. This can lead to complications
CC during transfusions or pregnancy if RhD negative individuals are exposed
CC to RhD positive blood, leading to them becoming immunised to produce
CC anti-D. The present invention relates to new human allo- and
CC auto-reactive T-cell epitopes (AAV99760-Y99769 and AAB03201-B03337) from
CC RhD, RhC, Rhc, RHe and RHe proteins. These epitopes bind to T-cells to
CC elicit an immune response, i.e. immunisation. These epitopes can be used
CC as a vaccine for the prevention of alloimmunisation or immunosuppression
CC of a response elicited by alloimmunisation or an autoimmune haemolytic
CC disease. Examples of autoimmune haemolytic diseases are sickle cell
CC disease and thalassaemia.
XX
SQ Sequence 15 AA;

Alignment Scores:
Pred. No.: 804 Length: 15
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 21 Gaps: 0

US-09-835-992A-22 (1-843) x AAB03322 (1-15)

OY 412 AAATTTTAAACCCCTA 429
Db 2 LyspheleuthrProleu 7
|||||

RESULT 12
ID AAV69422 standard; peptide: 25 AA.
XX
AC AAV69422;
XX
DT 03-JUL-2000 (first entry)
XX
DE Antimicrobial peptide derived from the protein Ranatuerin-1.
XX
KW Antibacterial; Rana peptide; Ranatuerin-1; Ranatuerin-2; Esculentin-2;
KW Brevinin-1; Ranalexin-1; Temporin-A; Temporin-B; American frog;
KW bacterial infection; Escherichia coli; Staphylococcus aureus;
KW Candida albicans; Enterococcus; Bacteroides fragilis;
KW Pseudomonas aeruginosa; Klebsiella pneumoniae; Serratia marcescens;
KW Mycobacterium tuberculosis; Streptococcus pneumoniae;
KW Streptococcus pyrogenes; Haemophilus influenzae;
KW Staphylococcus saprophyticus.
XX
OS Rana catesbeiana.
XX
PN WO200009553-A2.
XX
PD 24-FEB-2000.
XX
PF 13-AUG-1999; 99WO-US18575.
XX
PR 14-AUG-1998; 98US-0096607.
XX
PA (BION-) BIONEBRASKA INC.
XX (CONL/) CONLON J M.
XX
PI Conlon JM;
XX
DR WPI; 2000-237449/20.
XX
PT Antibacterial peptides selected from 7 different Rana peptide families
PT e.g. Ranatuerin-1, used to inhibit E. coli, S. aureus and C. albicans
PT infection -
XX
PS Claim 1; Page 35; 61pp; English.
XX
CC AAV69422-75 represent antibacterial peptides (termed Rana peptides)

CC which are derived from the Rana peptide families Ranatuerin-1,
CC Ranatuerin-2, Esculentin-2, Brevinin-1, Ranalexin-1, Temporin-A, or
CC Temporin-B. Each peptide is in the form of a C-terminus carboxylic
CC acid, a C-terminus amide or a C-terminus ester. The peptides are
CC derived from the skin of American frogs. The peptides can be used to
CC treat a bacterial infection in a patient. The peptides can be used
CC to inhibit Escherichia coli, Staphylococcus aureus, Candida albicans,
CC Enterococcus sp., Bacteroides fragilis, Pseudomonas aeruginosa,
CC Klebsiella pneumoniae, Serratia marcescens, Mycobacterium tuberculosis,
CC Streptococcus pneumoniae, Streptococcus pyrogenes, Haemophilus
CC influenzae, and staphylococcus saprophyticus.
XX
SQ Sequence 25 AA;

Alignment Scores:
Pred. No.: 744 Length: 25
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 21 Gaps: 0

US-09-835-992A-22 (1-843) x AAV69422 (1-25)

OY 96 TTAATAAACCTNGAATA 113
Db 6 LeuLysAsnLeuGlyLys 11
|||||

RESULT 13
ID AAV69423 standard; peptide: 25 AA.
XX
AC AAV69423;
XX
DT 03-JUL-2000 (first entry)
XX
DE Antimicrobial peptide derived from the protein Ranatuerin-1G.
XX
KW Antibacterial; Rana peptide; Ranatuerin-1; Ranatuerin-2; Esculentin-2;
KW Brevinin-1; Ranalexin-1; Temporin-A; Temporin-B; American frog;
KW bacterial infection; Escherichia coli; Staphylococcus aureus;
KW Candida albicans; Enterococcus; Bacteroides fragilis;
KW Pseudomonas aeruginosa; Klebsiella pneumoniae; Serratia marcescens;
KW Mycobacterium tuberculosis; Streptococcus pneumoniae;
KW Streptococcus pyrogenes; Haemophilus influenzae;
KW Staphylococcus saprophyticus; Ranatuerin-1G.
XX
OS Rana grylio.
XX
PN WO200009553-A2.
XX
PD 24-FEB-2000.
XX
PF 13-AUG-1999; 99WO-US18575.
XX
PR 14-AUG-1998; 98US-0096607.
XX
PA (BION-) BIONEBRASKA INC.
XX (CONL/) CONLON J M.
XX
PI Conlon JM;
XX
DR WPI; 2000-237449/20.
XX
PT Antibacterial peptides selected from 7 different Rana peptide families
PT e.g. Ranatuerin-1, used to inhibit E. coli, S. aureus and C. albicans
PT infection -
XX
PS Claim 1; Page 35; 61pp; English.
XX
CC AAV69422-75 represent antibacterial peptides (termed Rana peptides)
CC which are derived from the Rana peptide families Ranatuerin-1,
CC Ranatuerin-2, Esculentin-2, Brevinin-1, Ranalexin-1, Temporin-A, or

CC Temporin-B. Each peptide is in the form of a C-terminus carboxylic
CC acid, a C-terminus amide or a C-terminus ester. The peptides are
CC derived from the skin of American frogs. The peptides can be used
CC to treat a bacterial infection in a patient. The peptides can be used
CC to inhibit Escherichia coli, Staphylococcus aureus, Candida albicans,
CC Enterococcus sp., Bacteroides fragilis, Pseudomonas aeruginosa,
CC Klebsiella pneumoniae, Serratia marcescens, Mycobacterium tuberculosis,
CC Streptococcus pneumoniae, Streptococcus pyogenes, Haemophilus
CC influenzae, and Staphylococcus saprophyticus.

XX
SQ Sequence 25 AA;

Alignment Scores:
Pred. No.: 744 Length: 25
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 21 Gaps: 0

US-09-835-992A-22 (1-843) x AAY69423 (1-25)

OY 96 TTAACCTGCGGAA 113
DB 6 LeuLysAsnLeuGlyLys 11

RESULT 14
AAY69424
ID AAY69424 standard; peptide; 25 AA.
XX
AC AAY69424;
XX
DT 03-JUL-2000 (first entry)
XX
DE Antimicrobial peptide derived from the protein Ranatuerin-1C.
XX
KW Antibacterial; Rana peptide; Ranatuerin-1; Ranatuerin-2; Esculentin-2;
KW Brevinin-1; Ranalexin-1; Temporin-A; American frog;
KW bacterial infection; Escherichia coli; Staphylococcus aureus;
KW Candida albicans; Enterococcus; Bacteroides fragilis;
KW Pseudomonas aeruginosa; Klebsiella pneumoniae; Serratia marcescens;
KW Mycobacterium tuberculosis; Streptococcus pneumoniae;
KW Streptococcus pyogenes; Haemophilus influenzae;
KW Staphylococcus saprophyticus; Ranatuerin-1C.
XX
OS Rana clamitans.
XX
PN WO200009553-A2.
XX
PD 24-FEB-2000.
XX
PF 13-AUG-1999; 99WO-US18575.
XX
PR 14-AUG-1998; 98US-0096607.
XX
PA (BION-) BIONEERASKA INC.
PA (CONL/) CONLON J M.
XX
PI Conlon JM;
XX
DR WPI; 2000-237449/20.
XX
PT Antibacterial peptides selected from 7 different Rana peptide families
PT e.g. Ranatuerin-1, used to inhibit E. coli, S. aureus and C. albicans
PT infection -
XX
PS Claim 1; Page 35; 61pp; English.

CC derived from the skin of American frogs. The peptides can be used to
CC treat a bacterial infection in a patient. The peptides can be used
CC to inhibit Escherichia coli, Staphylococcus aureus, Candida albicans,
CC Enterococcus sp., Bacteroides fragilis, Pseudomonas aeruginosa,
CC Klebsiella pneumoniae, Serratia marcescens, Mycobacterium tuberculosis,
CC Streptococcus pneumoniae, Streptococcus pyogenes, Haemophilus
CC influenzae, and Staphylococcus saprophyticus.

XX
SQ Sequence 25 AA;

Alignment Scores:
Pred. No.: 744 Length: 25
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 21 Gaps: 0

US-09-835-992A-22 (1-843) x AAY69424 (1-25)

OY 96 TTAACCTGCGGAA 113
DB 6 LeuLysAsnLeuGlyLys 11

RESULT 15
AAY69469
ID AAY69469 standard; peptide; 25 AA.
XX
AC AAY69469;
XX
DT 03-JUL-2000 (first entry)
XX
DE Antimicrobial peptide derived from the protein Ranatuerin-1.
XX
KW Antibacterial; Rana peptide; Ranatuerin-1; Ranatuerin-2; Esculentin-2;
KW Brevinin-1; Ranalexin-1; Temporin-A; American frog;
KW bacterial infection; Escherichia coli; Staphylococcus aureus;
KW Candida albicans; Enterococcus; Bacteroides fragilis;
KW Pseudomonas aeruginosa; Klebsiella pneumoniae; Serratia marcescens;
KW Mycobacterium tuberculosis; Streptococcus pneumoniae;
KW Streptococcus pyogenes; Haemophilus influenzae;
KW Staphylococcus saprophyticus.
XX
OS Rana catesbeiana.
XX
FH Key Location/Qualifiers
FH
FT Misc-difference 3 /note= "not specified"
FT Misc-difference 16 /note= "not specified"
FT Misc-difference 21 /note= "not specified"
XX
PN WO200009553-A2.
XX
PD 24-FEB-2000.
XX
PF 13-AUG-1999; 99WO-US18575.
XX
PR 14-AUG-1998; 98US-0096607.
XX
PA (BION-) BIONEERASKA INC.
PA (CONL/) CONLON J M.
XX
PI Conlon JM;
XX
DR WPI; 2000-237449/20.
XX
PT Antibacterial peptides selected from 7 different Rana peptide families
PT e.g. Ranatuerin-1, used to inhibit E. coli, S. aureus and C. albicans
PT infection -
XX
PS Claim 2; Page 37; 61pp; English.

xx AAY69422-75 represent antibacterial peptides (termed Rana peptides)
CC which are derived from the Rana peptide families Ranatuerin-1,
CC Ranatuerin-2, Esculentin-2, Brevinin-1, Ranalexin-1, Temporin-A, or
CC Temporin-B. Each peptide is in the form of a C-terminus carboxylic
CC acid, a C-terminus amide or a C-terminus ester. The peptides are
CC derived from the skin of American frogs. The peptides can be used to
CC treat a bacterial infection in a patient. The peptides can be used
CC to inhibit Escherichia coli, Staphylococcus aureus, Candida albicans,
CC Enterococcus sp., Bacteroides fragilis, Pseudomonas aeruginosa,
CC Klebsiella pneumoniae, Serratia marcescens, Mycobacterium tuberculosis,
CC Streptococcus pneumoniae, Streptococcus pyogenes, Haemophilus
CC influenzae, and Staphylococcus saprophyticus.

xx
SQ Sequence 25 AA;

Alignment Scores:
Pred. NO.: 744 Length: 25
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 21 Gaps: 0

US-09-835-992a-22 (1-843) x AAY69469 (1-25)

OY 96 TTMAAAACCTNGAAA 113
Db 6 LeuLysAsnLeuGlyLys 11

Search completed: January 14, 2003, 17:27:28
Job time : 45.5252 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:54 ; Search time 11.5854 Seconds
(without alignments)
4281.862 Million cell updates/sec

Title: US-09-835-992A-22
Perfect score: 180
Sequence: 1 ggcacaaataattattttna.....taaatccgaacataaanaag 843

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+_n2p.model -DEV-olp
-O-/cgn2_1/USPRO-spool/US09835992/runat_14012003_161527_4050/app_query.fasta_1.3932
-DB-Issued_Patents_AA -QFMT-fastan -SUFFIX-olig.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.cdl
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE-LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09835992@cgn2_1_1_49@runat_14012003_161527_4050 -NCPUS=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6	3.3	35	2	US-08-142-551B-38	Sequence 38, Appl
2	6	3.3	147	4	US-09-325-932A-173	Sequence 173, App
3	6	3.3	164	3	US-08-765-381-2	Sequence 2, Appl
4	6	3.3	176	2	US-08-619-542B-44	Sequence 44, Appl
5	6	3.3	180	1	US-08-447-010-8	Sequence 8, Appl
6	6	3.3	187	4	US-08-981-392-46	Sequence 46, Appl
7	6	3.3	219	4	US-09-527-345-2	Sequence 2, Appl
8	6	3.3	247	4	US-09-414-828A-14	Sequence 14, Appl
9	6	3.3	257	1	US-08-451-947-10	Sequence 10, Appl
10	6	3.3	257	2	US-08-424-826A-10	Sequence 10, Appl
11	6	3.3	257	3	US-08-928-694-10	Sequence 10, Appl
12	6	3.3	257	5	PCT-US91-06950-10	Sequence 10, Appl

13	6	3.3	266	4	US-09-161-241-10	Sequence 10, Appl
14	6	3.3	273	2	US-08-701-935-5	Sequence 5, Appl
15	6	3.3	273	3	US-09-134-591-5	Sequence 5, Appl
16	6	3.3	291	4	US-09-420-786A-1	Sequence 1, Appl
17	6	3.3	295	3	US-08-461-607-13	Sequence 13, Appl
18	6	3.3	295	4	US-09-363-600-13	Sequence 13, Appl
19	6	3.3	313	2	US-08-990-379-7	Sequence 7, Appl
20	6	3.3	314	4	US-09-164-193-22	Sequence 22, Appl
21	6	3.3	314	4	US-09-221-448A-22	Sequence 22, Appl
22	6	3.3	319	3	US-09-100-664A-10	Sequence 10, Appl
23	6	3.3	319	4	US-09-335-983-10	Sequence 10, Appl
24	6	3.3	324	3	US-08-461-607-7	Sequence 7, Appl
25	6	3.3	324	4	US-09-363-600-7	Sequence 7, Appl
26	6	3.3	357	4	US-09-134-001C-3219	Sequence 3219, Ap
27	6	3.3	371	4	US-09-347-803-27	Sequence 27, Appl
28	6	3.3	390	2	US-08-873-093-3	Sequence 3, Appl
29	6	3.3	390	4	US-09-206-646-3	Sequence 3, Appl
30	6	3.3	398	4	US-08-856-841-21	Sequence 21, Appl
31	6	3.3	443	1	US-08-342-411A-4	Sequence 4, Appl
32	6	3.3	446	2	US-08-372-652-3	Sequence 3, Appl
33	6	3.3	446	4	US-08-776-844-2	Sequence 2, Appl
34	6	3.3	446	5	PCT-US95-16311-3	Sequence 3, Appl
35	6	3.3	460	1	US-08-342-411A-2	Sequence 2, Appl
36	6	3.3	461	1	US-08-330-518-2	Sequence 2, Appl
37	6	3.3	461	1	US-08-283-2	Sequence 2, Appl
38	6	3.3	461	2	US-08-646-248-2	Sequence 2, Appl
39	6	3.3	461	5	PCT-US95-13924-2	Sequence 2, Appl
40	6	3.3	461	5	PCT-US95-13931-2	Sequence 2, Appl
41	6	3.3	464	2	US-08-969-630-5	Sequence 5, Appl
42	6	3.3	470	1	US-08-514-014-2	Sequence 2, Appl
43	6	3.3	470	2	US-08-833-823-2	Sequence 2, Appl
44	6	3.3	486	4	US-09-134-001C-3593	Sequence 3593, Ap
45	6	3.3	504	4	US-09-126-420A-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-142-551B-38
: Sequence 38, Application US/08142551B
: Patent No. 5814603
: GENERAL INFORMATION:
: APPLICANT: Oldenburg, Kevin R.
: TITLE OF INVENTION: Selick, Harold E.
: TITLE OF INVENTION: COMPOUNDS WITH PTH ACTIVITY AND
: NUMBER OF SEQUENCES: 132
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Burns, Doane, Swecker & Mathis
: STREET: 699 Prince Street
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: US
: ZIP: 22313
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/142,551B
: FILING DATE: 25-OCT-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/077,296
: FILING DATE: 14-JUN-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/898,219
: FILING DATE: 12-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/965,677
: FILING DATE: 22-OCT-1992
: ATTORNEY/AGENT INFORMATION:

NAME: Swiss, Gerald F.
REGISTRATION NUMBER: 30,113
REFERENCE/DOCKET NUMBER: 000324-010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-7400
TELEFAX: (415) 854-8275
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 35
OTHER INFORMATION: /note="where "xaa" is selected
from the group consisting of Hol, Ho, a homoserine
OTHER INFORMATION: amide, or the sequence of amino acids comprising
OTHER INFORMATION: residues 35-84 of PTH."
US-08-142-551B-38

Alignment Scores:
Pred. No.: 227 Length: 35
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x US-08-142-551B-38 (1-35)

OY 96 TTAAACCTGGAAGA 113

Db 8 Leulysasnleucglylys 13

RESULT 2

US-09-325-932A-173

Sequence 173, Application US/09325932A

Patent No. 6451604

GENERAL INFORMATION:

APPLICANT: Flinn, Barry

APPLICANT: Lasham, Annette

TITLE OF INVENTION: Compositions affecting programmed cell

TITLE OF INVENTION: death and their use in the modification of forestry plant develop

FILE REFERENCE: 1022

CURRENT APPLICATION NUMBER: US/09/325,932A

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 206

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 173

LENGTH: 147

TYPE: PRT

ORGANISM: Pinus radiata

US-09-325-932A-173

Alignment Scores:

Pred. No.: 180

Score: 6.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 3.28%

DB: 4

Length: 147

Matches: 6

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-835-992A-22 (1-843) x US-09-325-932A-173 (1-147)

OY 627 CCGTTAATCCCAAGA 610

Db 65 Provallyssergingly 70

RESULT 3

US-08-765-381-2

Sequence 2, Application US/08765381

Patent No. 6083724

GENERAL INFORMATION:

APPLICANT: Commonwealth Scientific and Industrial Research Organisation

TITLE OF INVENTION: No. 6083724e1 avian cytokines and genetic

TITLE OF INVENTION: sequences encoding same

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully Scott Murphy and Presser

STREET: 400 Garden City Plaza

CITY: Garden City, New York

STATE: New York

COUNTRY: UNITED STATES OF AMERICA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/765,381

FILING DATE: 19-DEC-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU PN1542/95

FILING DATE: 06-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/AU96/00114

FILING DATE: 05-MAR-1996

ATTORNEY/AGENT INFORMATION:

NAME: PRESSER, LEOPOLD

TELECOMMUNICATION INFORMATION:

TELEPHONE: 1-516-742-4343

TELEFAX: 1-516-742-4366

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 164 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-765-381-2

Alignment Scores:

Pred. No.: 176

Score: 6.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 3.33%

DB: 3

Length: 164

Matches: 6

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-835-992A-22 (1-843) x US-08-765-381-2 (1-164)

OY 777 TTAAACACACCTTCC 794

Db 103 Leulysasnasnleupro 108

RESULT 4

US-08-619-542B-44

Sequence 44, Application US/08619542B

Patent No. 5830662

GENERAL INFORMATION:

APPLICANT: The Trustees of Columbia University in the City

APPLICANT: of New York

TITLE OF INVENTION: METHOD FOR CONSTRUCTION OF NORMALIZED

TITLE OF INVENTION: CDNA LIBRARIES

NUMBER OF SEQUENCES: 78

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/619,542B
FILING DATE: June 21, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 42840-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 176 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-619-542B-44

Alignment Scores:
Pred. No.: 174 Length: 176
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x US-08-619-542B-44 (1-176)
OY 767 GGTCCCCCCTTAAAAA 784
Db 110 G1yPp0r0l0eulyslys 115

RESULT 5
US-08-447-010-8
Sequence 8, Application US/08447010
Patent No. 5770718
GENERAL INFORMATION:
APPLICANT: MOFFATT, BARBARA
TITLE OF INVENTION: GENE FOR APRT FROM PLANT TISSUE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: 330 University Avenue, Suite 701
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,010
FILING DATE: 22-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/230,695
FILING DATE: 21-APR-1994
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,132
FILING DATE: 26-MAY-1992
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1811-183 MIS:V9
TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 180 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-447-010-8

Alignment Scores:
Pred. No.: 174 Length: 180
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-22 (1-843) x US-08-447-010-8 (1-180)
OY 771 GGACCNATCCCGTTTTC 754
Db 169 G1yPp0r0l0eProPh0phe 174

RESULT 6
US-08-981-392-46
Sequence 46, Application US/08981392
Patent No. 6262025
GENERAL INFORMATION:
APPLICANT: Ish-Horowicz, David
APPLICANT: Henrique, Domingos Manuel Pinto
APPLICANT: Lewis, Julian Hart
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Gray, Grace
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES
TITLE OF INVENTION: OF VERTEBRATE DELTA GENES AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/981,392
FILING DATE: 22-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Antler, Adriane M.
REGISTRATION NUMBER: 32,605
REFERENCE/DOCKET NUMBER: 7326-038
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-981-392-46

Alignment Scores:

Pred. No.: 173 Length: 187
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-22 (1-843) x US-08-981-392-46 (1-187)

OY 777 AGGGGGGACNATCCCG 760

Db 3 ArgglyglyProlePro 8

RESULT 7

US-09-527-345-2
; Sequence 2, Application US/09527345
; Patent No. 6331413
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Adler, David A.
; TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT APPLICATION NUMBER: US/09/527,345
; CURRENT FILING DATE: 1999-03-17
; PRIOR APPLICATION NUMBER: US 60/124,820
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-527-345-2

Alignment Scores:

Pred. No.: 168 Length: 219
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-527-345-2 (1-219)

OY 786 AACCTTCCCCNCCTTGTG 803

Db 46 AsnleuProProleu 51

RESULT 8

US-09-414-828A-14
; Sequence 14, Application US/09414828A
; Patent No. 6300084
; GENERAL INFORMATION:
; APPLICANT: Drubin, David G.
; APPLICANT: Hofmann, Christian J.
; TITLE OF INVENTION: Anti-Mitotic Agents and Processes
; FILE REFERENCE: Mitotic Spindle Proteins
; CURRENT APPLICATION NUMBER: US/09/414,828A
; CURRENT FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/103,684
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-414-828A-14

Alignment Scores:

Pred. No.: 165 Length: 247
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-414-828A-14 (1-247)

OY 777 TTAACCAACCTTCCC 794

Db 85 LeuLysAsnAsnleuPro 90

RESULT 9

US-08-451-947-10
; Sequence 10, Application US/08451947
; Patent No. 5702906
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; APPLICANT: ROSENTHAL, ARNON
; TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,947
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/426419
; FILING DATE: 19-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/030013
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/648482
; FILING DATE: 31-JAN
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/587707
; FILING DATE: 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: 666P2C1D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-451-947-10

Alignment Scores:

Pred. No.: 164 Length: 257
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-22 (1-843) x US-08-451-947-10 (1-257)

OY 378 ATTTTTCCTCCCAATCA 361

|||||
Db 26 llephepProasnpPro 31
RESULT 10
US-08-424-826A-10
; Sequence 10, Application US/08424826A
; Patent No. 5830858
; GENERAL INFORMATION:
; APPLICANT: Rosenthal, Arnon
; TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,826A
; FILING DATE: 19-Apr-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/240387
; FILING DATE: 10-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/648482
; FILING DATE: 31-Jan-1991
; APPLICATION NUMBER: 07/587707
; FILING DATE: 25-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Phd., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P0666P1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-424-826A-10
Alignment Scores:
Pred. No.: 164 Length: 257
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 2 Gaps: 0
US-09-835-992A-22 (1-843) x US-08-424-826A-10 (1-257)
QY 378 ATTTTTCCTCCCAATCCA 361
Db 26 llephepProasnpPro 31
RESULT 11
US-08-928-694-10
; Sequence 10, Application US/08928694
; Patent No. 6037320
; GENERAL INFORMATION:
; APPLICANT: ROSENTHAL, ARNON
; TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 100

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Genentech, Inc.
;; STREET: 1 DNA Way
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WinPatIn (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/928,694
;; FILING DATE: 12-Sep-1997
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/451947
;; FILING DATE: 26-MAY-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/426419
;; FILING DATE: 19-APR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/030013
;; FILING DATE: 22-MAR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/648482
;; FILING DATE: 31-JAN
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/587707
;; FILING DATE: 1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Torchia, Phd., Timothy E.
;; REGISTRATION NUMBER: 36,700
;; REFERENCE/DOCKET NUMBER: P0666P2C1D2C1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650/225-8674
;; TELEFAX: 650/952-9881
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 257 amino acids
;; TYPE: Amino Acid
;; TOPOLOGY: Linear
;; US-08-928-694-10
Alignment Scores:
Pred. No.: 164 Length: 257
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 3 Gaps: 0
US-09-835-992A-22 (1-843) x US-08-928-694-10 (1-257)
QY 378 ATTTTTCCTCCCAATCCA 361
Db 26 llephepProasnpPro 31
RESULT 12
PCT-US91-06950-10
; Sequence 10, Application PC/TUS9106950
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; APPLICANT: ROSENTHAL, ARNON
; TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA

ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 Inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/06950
FILING DATE: 19910924
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648482
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/587707
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,043
REFERENCE/DOCKET NUMBER: 666P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-06950-10

Alignment Scores:
Pred. No.: 164 Length: 257
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 5 Gaps: 0

US-09-835-992a-22 (1-843) x PCT-US91-06950-10 (1-257)

OY 378 ATTTTTCCTCCCAATCCA 361
Db 26 lIephepHeProasnpRo 31

RESULT 13
US-09-161-241-10

; Sequence 10, Application US/09161241
; Patent No. 6344541

; GENERAL INFORMATION:

; APPLICANT: Bass, Michael B

; APPLICANT: Sullivan, John K

; APPLICANT: Theill, Lars E

; APPLICANT: Wang, Daquang

; TITLE OF INVENTION: NOVEL DKR POLYPEPTIDES

; FILE REFERENCE: A-548

; CURRENT APPLICATION NUMBER: US/09/161,241

; CURRENT FILING DATE: 1998-09-25

; NUMBER OF SEQ ID NOS: 78

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10

; LENGTH: 266

; TYPE: PRT

; ORGANISM: Human

; US-09-161-241-10

Alignment Scores:
Pred. No.: 163 Length: 266
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 4 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-161-241-10 (1-266)

OY 786 AACCTTCCCCNCCTTG 803
Db 44 AsnleuProProleu 49

RESULT 14
US-08-701-935-5

; Sequence 5, Application US/08701935
; Patent No. 5843683

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Au-Young, Janice

; APPLICANT: Hillman, Jennifer

; APPLICANT: Goll, Surya K.

; TITLE OF INVENTION: NOVEL HUMAN PROTEASOME

; TITLE OF INVENTION: SUBUNIT PROTEINS

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/701,935

; FILING DATE: Filed Herewith

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0119 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 273 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 307307

; US-08-701-935-5

Alignment Scores:

Pred. No.: 162 Length: 273

Score: 6.00 Matches: 6

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.33% Indels: 0

DB: 2 Gaps: 0

US-09-835-992a-22 (1-843) x US-08-701-935-5 (1-273)

OY 617 GATTTAACGGGTCCCAA 634
Db 143 AspleuThrglyProglIn 148

RESULT 15
US-09-134-591-5

; Sequence 5, Application US/09134591
; Patent No. 6018028

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Au-Young, Janice

; APPLICANT: Hillman, Jennifer

APPLICANT: COLL, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PROTEASOME
TITLE OF INVENTION: SUBUNIT PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,591
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/701,935
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0119 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 307307
US-09-134-591-5

Alignment Scores:
pred. No.: 162 Length: 273
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 3 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-134-591-5 (1-273)

OY 617 GATTTAACGGGTCCCAA 634
Db 143 AspleuthrGlyProGln 148

Search completed: January 14, 2003, 17:42:19
Job time : 14.8354 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:29:25 ; Search time 8.07072 Seconds
(without alignments)
4052.943 Million cell updates/sec

Title: US-09-835-992A-22
Perfect score: 180
Sequence: 1 ggccaaaaaantattattna.....taaatcgaacaataaaaag 843

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 118974 seqs, 19401057 residues

Word size: 1

Total number of hits satisfying chosen parameters: 214280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US9835992/runat_14012003_161528_4118/app_query.fasta_1.3932
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=olig.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09835992@cgn_1.1.28 @runat_14012003_161528_4118
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	6	3.3	46	10	US-09-864-761-47474	Sequence 47474, A
C 2	6	3.3	74	10	US-09-764-847-677	Sequence 677, App
C 3	6	3.3	91	10	US-09-764-877-1669	Sequence 1669, Ap
C 4	6	3.3	129	10	US-09-867-550-1822	Sequence 1822, Ap

5	6	3.3	140	10	US-09-893-737-118	Sequence 118, App
6	6	3.3	140	10	US-09-746-491-6	Sequence 6, Appl1
7	6	3.3	182	10	US-09-764-903-47	Sequence 47, Appl
C 8	6	3.3	187	10	US-09-908-322-46	Sequence 46, Appl
9	6	3.3	191	9	US-09-738-626-5878	Sequence 5878, Ap
10	6	3.3	219	9	US-09-922-469-2	Sequence 2, Appl1
11	6	3.3	219	10	US-09-922-480-2	Sequence 2, Appl1
C 12	6	3.3	219	10	US-09-923-236-2	Sequence 10, Appl1
C 13	6	3.3	257	8	US-08-450-842-10	Sequence 10, Appl
C 14	6	3.3	259	9	US-09-996-634-133	Sequence 133, App
15	6	3.3	266	9	US-09-974-298-89	Sequence 89, Appl
16	6	3.3	266	9	US-09-978-295A-456	Sequence 456, App
17	6	3.3	266	9	US-09-976-736-10	Sequence 456, App
18	6	3.3	266	9	US-09-978-697-456	Sequence 456, App
19	6	3.3	266	9	US-09-978-192A-456	Sequence 456, App
20	6	3.3	266	9	US-09-999-832A-456	Sequence 456, App
21	6	3.3	266	9	US-09-978-189-456	Sequence 456, App
22	6	3.3	266	9	US-10-028-072-428	Sequence 428, App
C 23	6	3.3	297	10	US-09-925-297-635	Sequence 635, App
C 24	6	3.3	329	9	US-09-895-913A-170	Sequence 170, App
25	6	3.3	361	9	US-09-738-626-6663	Sequence 6663, Ap
26	6	3.3	373	9	US-10-137-731-2	Sequence 2, Appl1
27	6	3.3	390	9	US-10-072-130-3	Sequence 3, Appl1
28	6	3.3	409	10	US-09-790-264-42	Sequence 42, Appl
C 29	6	3.3	437	10	US-09-790-264-30	Sequence 30, Appl
C 30	6	3.3	446	10	US-09-909-446-2	Sequence 2, Appl1
C 31	6	3.3	446	10	US-09-909-325-2	Sequence 2, Appl1
C 32	6	3.3	446	10	US-09-909-326-2	Sequence 2, Appl1
C 33	6	3.3	446	12	US-10-013-823-2	Sequence 2, Appl1
C 34	6	3.3	458	9	US-09-738-626-3686	Sequence 3686, Ap
C 35	6	3.3	461	12	US-10-013-823-3	Sequence 3, Appl1
36	6	3.3	470	9	US-10-114-893-8	Sequence 8, Appl1
37	6	3.3	480	10	US-09-790-264-32	Sequence 32, Appl
38	6	3.3	482	10	US-09-796-858-40	Sequence 40, Appl
39	6	3.3	593	10	US-09-836-410A-1	Sequence 1, Appl1
C 40	6	3.3	593	10	US-09-836-503B-2	Sequence 2, Appl1
C 41	6	3.3	630	10	US-09-843-598-10	Sequence 10, Appl
42	6	3.3	700	10	US-09-836-503B-5	Sequence 5, Appl1
43	6	3.3	710	10	US-09-836-503B-4	Sequence 4, Appl1
44	6	3.3	723	10	US-09-836-503B-3	Sequence 3, Appl1
C 45	6	3.3	748	10	US-09-864-761-43244	Sequence 43244, A

ALIGNMENTS

RESULT 1
US-09-864-761-47474
Sequence 47474, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47474
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC022325.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EST_HUMAN HIT: AV681548.1, EVALU 5.20e+00
; OTHER INFORMATION: SWISSPROT HIT: O76039, EVALU 8.20e+00
US-09-864-761-47474

Alignment Scores:
Pred. No.: 161 Length: 46
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-864-761-47474 (1-46)
QY 802 AAAGNGGGGGAAGTTG 785
Db 39 LysGlyGlyGlyArgLeu 44

RESULT 2
US-09-764-847-677
; Sequence 677, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 677
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (33)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (53)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-677

Alignment Scores:
Pred. No.: 145 Length: 74
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-764-847-677 (1-74)
QY 376 TTTTCCCAATCCAA 359
Db 61 PhepheProGlnSerLys 66

RESULT 3
US-09-764-877-1669
; Sequence 1669, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1669
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (15)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1669

Alignment Scores:
Pred. No.: 139 Length: 91
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-764-877-1669 (1-91)
QY 777 AGGGGGGACCNATCCG 760
Db 61 ArgGlyGlyProIlePro 66

RESULT 4
US-09-867-550-1822
; Sequence 1822, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehtraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)

; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1822
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (72)
; OTHER INFORMATION: wherein xaa may be any one of Arg or Cys or Gly or Ser
; NAME/KEY: VARIANT
; LOCATION: (92)
; OTHER INFORMATION: wherein xaa may be any one of Gln or His
US-09-867-550-1822

Alignment Scores:
Pred. No.: 129 Length: 129
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-867-550-1822 (1-129)

OY 635 ATTGGGACCCCTTAAAT 618
Db 121 IleglyasProLeuAa 126

RESULT 5

; US-09-893-737-118
; Sequence 118, Application US/09893737
; Patent No. US2002011085A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 118
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-118

Alignment Scores:
Pred. No.: 127 Length: 140
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-893-737-118 (1-140)

OY 789 CTTCCCCNCCTTGGCC 806
Db 11 LeuProProLeuAa 16

RESULT 6

; US-09-746-491-6
; Sequence 6, Application US/09746491
; Patent No. US20020137202A1
; GENERAL INFORMATION:

; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20020137202A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-621
; CURRENT APPLICATION NUMBER: US/09/746,491
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: USSN 60/171,329
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-746-491-6

Alignment Scores:
Pred. No.: 127 Length: 140
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-746-491-6 (1-140)

OY 789 CTTCCCCNCCTTGGCC 806
Db 11 LeuProProLeuAa 16

RESULT 7

; US-09-764-903-47
; Sequence 47, Application US/09764903
; Patent No. US20020090674A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ28
; CURRENT APPLICATION NUMBER: US/09/764,903
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-903-47

Alignment Scores:
Pred. No.: 120 Length: 182
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-764-903-47 (1-182)

OY 792 CCCCNCCTTGGCCAG 809
Db 123 ProProProLeuAa 128

RESULT 8

; US-09-908-322-46
; Sequence 46, Application US/09908322
; Patent No. US20020107194A1
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; Henriques, Domingos Manuel Pinto
; Lewis, Julian Hart
; Artavanis-Tsakonas, Spyridon
; Gray, Grace

; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF

VERTEBRATE DELTA GENE AND METHODS BASED THEREON

NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036/2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/908,322
FILING DATE: 17-Jul-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/981,392
FILING DATE: 22-DEC-1997

ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-123

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-908-322-46

Alignment Scores:
Pred. No.: 119 Length: 187
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-908-322-46 (1-187)

QY 777 AGGGGGGACCNATCCCG 760
|||||
Db 3 ArgGlyGlyProIlePro 8

RESULT 9
US-09-738-626-5878
Sequence 5878, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484

PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5878
LENGTH: 191
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5878

Alignment Scores:
Pred. No.: 119 Length: 191
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 9 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-738-626-5878 (1-191)

QY 616 GGATTTACGGGTCCTCCA 633
|||||
Db 72 GlyPheAsnGlySerPro 77

RESULT 10
US-09-922-469-2
Sequence 2, Application US/09922469
Patent No. US20020173027A1
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
FILE REFERENCE: 97-71
CURRENT APPLICATION NUMBER: US/09/922,469
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US 60/124,820
PRIOR FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 219
TYPE: PRT
ORGANISM: Homo sapiens
US-09-922-469-2

Alignment Scores:
Pred. No.: 115 Length: 219
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 9 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-922-469-2 (1-219)

QY 786 AACCTTCCCCNCCTTGG 803
|||||
Db 46 AsnLeuProProIleu 51

RESULT 11
US-09-922-480-2
Sequence 2, Application US/09922480
Patent No. US20020081701A1
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Adler, David A.
TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
FILE REFERENCE: 97-71
CURRENT APPLICATION NUMBER: US/09/922,480
CURRENT FILING DATE: 2001-08-03

;; PRIOR APPLICATION NUMBER: US 60/124,820
;; PRIOR FILING DATE: 1999-03-17
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-480-2

Alignment Scores:
Pred. No.: 115 Length: 219
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-922-480-2 (1-219)
QY 786 AACCTTCCCCNCCTTTG 803
DB 46 AsnleuProProProleu 51

RESULT 12
US-09-923-236-2
; Sequence 2, Application US/09923236
; Patent No. US20020090677A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Adler, David A.
; TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT APPLICATION NUMBER: US/09/923,236
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/124,820
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-236-2

Alignment Scores:
Pred. No.: 115 Length: 219
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-923-236-2 (1-219)
QY 786 AACCTTCCCCNCCTTTG 803
DB 46 AsnleuProProProleu 51

RESULT 13
US-08-450-842-10
; Sequence 10, Application US/08450842
; Patent No. US20020045576A1
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; APPLICANT: ROSENTHAL, ARNON
; TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California

;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: patln (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,842
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/426419
; FILING DATE: 19-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/030013
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/648482
; FILING DATE: 31-JAN
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/587707
; FILING DATE: 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: 666P2C1D3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/952-9881
; TELEFAX: 415/225-8674
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-450-842-10

Alignment Scores:
Pred. No.: 111 Length: 257
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 8 Gaps: 0

US-09-835-992A-22 (1-843) x US-08-450-842-10 (1-257)
QY 378 ATTTTTCCTCCCAATCCA 361
DB 26 IlePhePheProAsnPro 31

RESULT 14
US-09-996-634-133
; Sequence 133, Application US/09996634
; Patent No. US20020172684A1
; GENERAL INFORMATION:
; APPLICANT: Nano, Francis
; TITLE OF INVENTION: Mycobacterium Tuberculosis DNA Sequences Encoding
; FILE REFERENCE: 61260
; CURRENT APPLICATION NUMBER: US/09/996,634
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/447,135
; PRIOR FILING DATE: 2000-01-03
; PRIOR APPLICATION NUMBER: 08/990,823
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 96/10375
; PRIOR FILING DATE: 1996-06-14
; PRIOR APPLICATION NUMBER: 60/000,254
; PRIOR FILING DATE: 1995-06-15
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: Patentln Ver. 2.0

; SEQ ID NO 133
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-996-634-133

Alignment Scores:
Pred. No.: 111 Length: 259
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 9 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-996-634-133 (1-259)

OY 802 AAAGNGGGGAGGTTG 785
|||||
DB 204 LysGlyGlyGlyArgLeu 209

RESULT 15

US-09-974-298-89
; Sequence 89, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 89
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 2481150CD1
US-09-974-298-89

Alignment Scores:
Pred. No.: 111 Length: 266
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 9 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-974-298-89 (1-266)

OY 786 AACCTTCCCGCNCCTTTG 803
|||||
DB 44 AsnLeuProProLeu 49

Search completed: January 14, 2003, 18:08:22
Job time : 10.0707 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:45 ; Search time 28.2475 Seconds
(without alignments)
5737.940 Million cell updates/sec

Title: US-09-835-992A-22
Perfect score: 180
Sequence: 1 ggcccaaaaantattctna.....taaatccgaacataaaaaag 843

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565536

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+_n2p.model -DEV-xlp
-O-/cgn2.1/USPTO_spool/US09835992/runat_14012003_161526_4038/app-query.fasta_1.3932
-DB-PIR_73 -QFMT-fastan -SUFFIX-olig.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09835992_@CGN_1.1.192_@runat_14012003_161526_4038 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	8	4.4	497	2 T21579	hypothetical prote
C 2	7	3.8	86	2 S44154	hypothetical prote
C 3	7	3.9	172	2 A70864	hypothetical prote
4	7	3.9	223	2 G90126	hypothetical prote
5	7	3.9	337	2 A81255	DNA-directed RNA p
6	7	3.9	387	2 T13114	sopa protein homol
7	7	3.9	435	2 S68865	fimbrial protein h
8	7	3.9	498	2 T47919	hypothetical prote
9	7	3.9	1240	2 T48800	SMT4 related prote
10	7	3.9	1371	1 VCBEM7	major capsid prote
11	7	3.9	1371	2 T42938	major capsid prote
C 12	6	3.3	60	2 T07328	hypothetical prote
C 13	6	3.3	68	2 F81945	probable hemeolysi
14	6	3.3	90	2 C83058	hypothetical prote

C 15	6	3.3	90	2 T27091	hypothetical prote
C 16	6	3.3	92	2 A71375	hypothetical prote
17	6	3.3	94	2 S62084	M-like protein enn
18	6	3.3	95	2 A86801	prophage p13 prote
19	6	3.3	106	2 S61176	hypothetical prote
C 20	6	3.3	110	2 T48382	hypothetical prote
C 21	6	3.3	117	2 S47178	hypothetical prote
C 22	6	3.3	119	2 T48745	hypothetical prote
C 23	6	3.3	126	2 S58321	probable membrane
C 24	6	3.3	127	2 F90096	high light inducib
C 25	6	3.3	156	2 T44769	hypothetical prote
C 26	6	3.3	161	2 S76047	hypothetical prote
27	6	3.3	162	2 B72534	hypothetical prote
C 28	6	3.3	169	2 G86609	L9 ribosomal prote
29	6	3.3	169	2 H72014	ribosomal protein
30	6	3.3	169	2 E69362	conserved hypothet
31	6	3.3	173	2 E81258	probable membrane
C 32	6	3.3	174	2 AE2528	hypothetical prote
C 33	6	3.3	177	2 A86735	transcription regu
C 34	6	3.3	180	1 RTMSA	adenine phosphorib
C 35	6	3.3	180	2 E97200	probable phosphata
36	6	3.3	183	2 C75386	RNA-directed cyclin
C 37	6	3.3	184	2 S65817	hypothetical prote
C 38	6	3.3	185	2 A13454	hypothetical prote
39	6	3.3	185	2 F87598	hypothetical prote
40	6	3.3	186	2 B97353	hypothetical prote
41	6	3.3	190	2 T19826	ADP-ribosylation f
42	6	3.3	205	2 I40179	hypothetical prote
43	6	3.3	207	2 C69780	hypothetical prote
44	6	3.3	208	2 AC3091	hypothetical prote
C 45	6	3.3	214	2 C95195	hypothetical prote

ALIGNMENTS

RESULT 1
T21579
hypothetical protein F30F8.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T21579
R:Wilkinson, J.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19443
A:Accession: T21579
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-497 <WIL>
A:Cross-references: EMBL:281073; PIDN:CAB03034.1; GSPDB:GN00019; CESP:F30F8.3
A:Experimental source: clone F30F8
C:Genetics:
A:Gene: CESP:F30F8.3
A:Map position: 1
A:Introns: 36/3; 109/1; 189/1; 235/2; 276/1; 314/3; 342/2; 384/3; 404/3; 451/3; 479/1
C:Superfamily: Caenorhabditis elegans hypothetical protein F30F8.3

Alignment Scores:
Pred. No.: 1.42
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 4.37%
DB: 2
Gaps: 0

US-09-835-992A-22 (1-843) x T21579 (1-497)

OY 127 TTTGCNATATTTTTCNAGG 104
Db 340 PheAlaIlellePhePheProArg 347

RESULT 2
S44154
hypothetical protein 1 - wheat

C:Species: Triticum aestivum (common wheat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-Oct-1999
C:Accession: S44154
R:Kroczyńska, B.; Buchowicz, J.
submitted to the EMBL Data Library, November 1993
A:Reference number: S44153
A:Accession: S44154
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <KRO>
A:Cross-references: EMBL:X76181; NID:g473172; PIDN:CAA53774.1; PID:g473173

Alignment Scores:
Pred. No.: 27.9 Length: 86
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.83% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x S44154 (1-86)

OY 799 GGNGCGGAGCTGTTT 779
|||||
Db 37 G|yG|yG|yA|rgLeuPhephe 43

RESULT 3
A70864
hypothetical protein Rv2450c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: A70864
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70864
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-172 <COL>
A:Cross-references: GB:AL021246; GB:AL123456; NID:g3261507; PIDN:CAA16027.1; PID:el23757
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv2450c

Alignment Scores:
Pred. No.: 24.7 Length: 172
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x A70864 (1-172)

OY 786 AACCTGCCCNCTTTGGCC 806
|||||
Db 73 AsnLeuProProLeuAla 79

RESULT 4
G90126
hypothetical protein orf223 [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: G90126
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reijnders, L.J.
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671; PMID:11323671

A:Accession: G90126
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-223 <DOU>
A:Cross-references: GB:AF083031; NID:g13794330; PIDN:AAK39707.1; GSPDB:GN00152
C:Genetics:
A:Gene: orf223
A:Map position: 3
A:Genome: nucleomorph
C:Keywords: nucleomorph

Alignment Scores:
Pred. No.: 23.5 Length: 223
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x G90126 (1-223)

OY 99 AAAAACCCTNGGAAAAAATA 119
|||||
Db 87 LysAsnLeuGlyLysLysIle 93

RESULT 5
A81255
DNA-directed RNA polymerase (EC 2.7.7.6) alpha chain Cj1595 [imported] - Campylobacter jejuni
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: A81255
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chl C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Bar Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: A81255
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <PAR>
A:Cross-references: GB:AL139079; GB:AL111168; NID:g6968971; PIDN:CAB73583.1; PID:g696
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: rpoA; Cj1595
C:Superfamily: DNA-directed RNA polymerase alpha chain
C:Keywords: nucleotidyltransferase

Alignment Scores:
Pred. No.: 21.8 Length: 337
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x A81255 (1-337)

OY 96 TTAAAAAACCTNGGAAAAA 116
|||||
Db 290 LeuLysAsnLeuGlyLysLys 296

RESULT 6
T13114
sopa protein homolog - phage N15
N:Alternate names: protein gp28
C:Species: phage N15
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 20-Sep-1999
C:Accession: T13114
R:Hendrix, R.W.; Ravin, V.K.; Casjens, S.R.; Ford, M.E.; Ravin, N.V.; Smirnov, I.K.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z17603
A:Accession: T13114
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-387 <HEN>
A:Cross-references: EMBL:AF064539; NID:g3192683; PID:g3192711; PIDN:AAC19064.1
C:Genetics:
A:Note: gene 28
C:Superfamily: sopa protein

Alignment Scores:

Pred. No.:	21.3	Length:	387
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.89%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-22 (1-843) x T13114 (1-387)

QY 670 TTACCAATTTTACCTTGCTAAG 690
|||||
Db 169 LeuProPheTyrLeuGlyLys 175

RESULT 7

S68865
fimbrial protein h1fe precursor - Haemophilus influenzae
C:Species: Haemophilus influenzae
C:Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 08-Oct-1999
C:Accession: S68865; S54431
R:Van Ham, M.S.
submitted to the EMBL Data Library, May 1994
A:Reference number: S68865
A:Accession: S68865
A:Molecule type: DNA
A:Residues: 1-435 <VAN>
A:Cross-references: EMBL:Z33502; NID:g535165; PIDN:CAA83904.1; PID:g535169
R:Van Ham, S.M.; van Alphen, L.; Mool, F.R.; van Putten, J.P.M.
Mol. Microbiol. 13, 673-684, 1994
A:Title: The fimbrial gene cluster of Haemophilus influenzae type b.
A:Reference number: S54428; MUID:95089703; PMID:7997179
A:Accession: S54431
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 406-435 <VAV>
A:Cross-references: EMBL:Z33502
C:Genetics:
A:Gene: h1fe

Alignment Scores:

Pred. No.:	20.9	Length:	435
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.89%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-22 (1-843) x S68865 (1-435)

QY 776 CTTAAACACACCTTCCCC 796
|||||
Db 389 LeuLysGlnProSerPro 395

RESULT 8

T47919
hypothetical protein T20K12.170 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Jun-2000
C:Accession: T47919
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quent
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z24480
A:Accession: T47919
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-498 <DEH>

A:Cross-references: EMBL:AL137898
A:Experimental source: cultivar Columbia; BAC clone T20K12
C:Genetics:
A:Map position: 3
A:Introns: 46/3; 222/2; 286/3; 308/2; 378/3
A:Note: T20K12.170
C:Superfamily: Arabidopsis thaliana hypothetical protein T20K12.190

Alignment Scores:

Pred. No.:	20.4	Length:	498
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.89%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-22 (1-843) x T47919 (1-498)

QY 817 TCCCGTCTAATCCGAACAT 837
|||||
Db 84 SerArgLeuAsnProAsnAsn 90

RESULT 9

T48800
SMT4 related protein [imported] - Neurospora crassa
N:Alternate names: protein 15E6.80
C:Species: Neurospora crassa
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C:Accession: T48800
R:Schulte, U.; Algn, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24541
A:Accession: T48800
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1240 <SCH>
A:Cross-references: EMBL:AL353822; GSPDB:GN00112; NCSP:15E6.80
A:Experimental source: cosmid contig 15E6; strain 74
C:Genetics:
A:Gene: NCSP:15E6.80
A:Map position: 2
A:Introns: 8/3; 358/2

Alignment Scores:

Pred. No.:	17.3	Length:	1240
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.89%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-22 (1-843) x T48800 (1-1240)

QY 767 GGTCCCCCCTTAAAAACAA 787
|||||
Db 29 GlyProProLeuLysLysGln 35

RESULT 10

VCBEM7
major capsid protein - salmirline herpesvirus 1 (strain 11)
C:Species: salmirline herpesvirus 1
A:Note: host Salmirl sciureus (common squirrel monkey)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999
C:Accession: E36808
R:Albrecht, J.
submitted to the EMBL Data Library, January 1992
A:Description: Primary structure of the herpesvirus salmirl genome.
A:Reference number: A36806
A:Accession: E36808
A:Molecule type: DNA
A:Residues: 1-1371 <ALB>
A:Cross-references: GB:X64346; NID:g60320; PIDN:CAA45648.1; PID:g60346
R:Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.;

J. Virol. 66, 5047-5058, 1992

A;Title: Primary structure of the herpesvirus saimiri genome.

A;Reference number: A37309; MUID:9233688; PMID:1321287

A;Contents: annotation; possible protein-coding frames

A;Note: neither amino acid nor nucleotide sequence is given

C;Genetics:

A;Gene: 25

C;Superfamily: varicella-zoster virus major capsid protein

C;Keywords: capsid protein

Alignment Scores:

Pred. No.:	17	Length:	1371
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.89%	Indels:	0
DB:	1	Gaps:	0

US-09-835-992A-22 (1-843) x VCBEM7 (1-1371)

QY 786 AACCTTCCCCNCCTTGGCC 806

|||||

Db 560 AsnLeuProProLeuAla 566

RESULT 11

major capsid protein - ateline herpesvirus 3 (strain 73)

C;Species: ateline herpesvirus 3

A;Variety: strain 73

C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C;Accession: T42938

R;Albrecht, J.C.; Fleckenstein, B.

submitted to the EMBL Data Library, August 1998

A;Description: Primary structure of the herpesvirus ateles genome.

A;Reference number: Z22274

A;Accession: T42938

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1371 <ALB>

A;Cross-references: EMBL:AF083424; PIDN:AAC95549.1

A;Experimental source: strain 73

C;Superfamily: varicella-zoster virus major capsid protein

Alignment Scores:

Pred. No.:	17	Length:	1371
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.89%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-22 (1-843) x T42938 (1-1371)

QY 786 AACCTTCCCCNCCTTGGCC 806

|||||

Db 560 AsnLeuProProLeuAla 566

RESULT 12

T07328

hypothetical protein 60c - Chlorella vulgaris chloroplast

C;Species: chloroplast Chlorella vulgaris

C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000

C;Accession: T07328

R;Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Tsudzuki, J.; Nakas

Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997

A;Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlo

A;Reference number: Z15985; MUID:97303241; PMID:9159184

A;Accession: T07328

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-60 <WAK>

A;Cross-references: EMBL:AB001684; NID:g22224352; PIDN:BAA57976.1; PID:g22224492

C;Genetics:

A;Genome: chloroplast

C;Keywords: chloroplast

Alignment Scores:

Pred. No.:	426	Length:	60
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.28%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-22 (1-843) x T07328 (1-60)

QY 784 TTTTAAAGGGGGGACC 767

|||||

Db 28 PhePhelysGlyGlyThr 33

RESULT 13

F81945

probable hemolysin MMA0983 [imported] - Neisseria meningitidis (strain Z2491 serogro

C;Species: Neisseria meningitidis

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C;Accession: F81945

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: F81945

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-68 <PAR>

A;Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84253.1; PID:g737

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: MMA0983

Alignment Scores:

Pred. No.:	417	Length:	68
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.28%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-22 (1-843) x F81945 (1-68)

QY 805 GCCAAGGNGGGGAGG 788

|||||

Db 31 AlaLysGlyGlyArg 36

RESULT 14

C83058

hypothetical protein PA4703 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C;Accession: C83058

R;Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: C83058

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-90 <STO>

A;Cross-references: GB:AE004884; GB:AE004091; NID:g9950954; PIDN:AAG08089.1; GSPDB:GN

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA4703

Alignment Scores:

Pred. No.:	396	Length:	90
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.33%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-22 (1-843) x C83058 (1-90)

QY	96	TTAAAAAACCCTNGAANA	113
Db	8	LeuLysAsnLeuClLys	13

RESULT 15

hypothetical protein Y51B9A.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27091
R:Barlow, K.
submitted to the EMBL Data Library, June 1998
A:Reference number: Z20308
A:Accession: T27091
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-90 <WIL>
A:Cross-references: EMBL:AL023845; PIDN:CAA19540.1; GSPDB:GN00020; CESP:Y51B9A.5
A:Experimental source: clone Y51B9A
C:Genetics:
A:Gene: CESP:Y51B9A.5
A:Map position: 2
A:Introns: 21/3

Alignment Scores:

Pred. No.:	396	Length:	90
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.33%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-22 (1-843) x T27091 (1-90)

QY 792 CCCCCNCCTTTGGCCAC 809
 |||
 Db 45 PROPROLEUALAGln 50

Search completed: January 14, 2003, 17:40:51
Job time : 34.2475 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:54 ; Search time 13.0173 Seconds
(without alignments)
5372.014 Million cell updates/sec

Title: US-09-835-992A-22
Perfect score: 180
Sequence: 1 ggccaaaaaaantattctna.....taaatccgaacaataaaaag 843

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+n2p.model -DEV-xlp
-Q=/cgn2_1/USPTO_spool/US09835992/runat_14012003_161525_4014/app-query.fasta_1.3932
-DB=SwissProt_40 -QFMT=fastn -SUFFIX=olig.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS-human40.cdl
-LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992@cgn_1_1_77@runat_14012003_161525_4014 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	3.9	337	1	RPOA_CAMJF
2	7	3.9	435	1	HFE1_HAEIN
3	7	3.9	1371	1	VCAP_HSVSA
4	6	3.3	92	1	Y031_TREPA
5	6	3.3	164	1	ING_CHICK
6	6	3.3	164	1	ING_MELGA
7	6	3.3	164	1	ING_PHACO
8	6	3.3	169	1	RL9_CHLPN
9	6	3.3	180	1	APT_MASHI
10	6	3.3	180	1	APT_MOUSE
11	6	3.3	180	1	APT_MUSPA
12	6	3.3	180	1	APT_MUSSI
13	6	3.3	183	1	RRE_DEIRA
14	6	3.3	218	1	VPDC_BACSU
15	6	3.3	226	1	YRAL_YEAST
16	6	3.3	227	1	TPIS_SULSO
17	6	3.3	238	1	TCBE_PSESO
18	6	3.3	239	1	PNUC_SALTY

C	19	6	3.3	241	1	PSMA_SULSO	O9uxc6	sulfolobus
C	20	6	3.3	247	1	YGG1_YEAST	P53168	saccharomyc
C	21	6	3.3	254	1	HIS6_LEPBO	O9zgm0	leptosira
C	22	6	3.3	254	1	HIS6_LEPIN	O9s4h7	leptosira
C	23	6	3.3	257	1	NT6B_HUMAN	P34133	homo sapien
C	24	6	3.3	261	1	LPRE_MYCTU	P71798	mycobacteri
C	25	6	3.3	263	1	DKK1_SPIAP	P35892	splroplasma
C	26	6	3.3	266	1	DKK1_HUMAN	O94907	homo sapien
C	27	6	3.3	266	1	PR11_YEAST	O07350	saccharomyc
C	28	6	3.3	269	1	HXA9_MORSA	O9pws5	morone saxa
C	29	6	3.3	269	1	THID_TREPA	O83153	treponema p
C	30	6	3.3	272	1	DKK1_MOUSE	O54908	mus musculu
C	31	6	3.3	273	1	PSBA_HUMAN	P40306	homo sapien
C	32	6	3.3	284	1	RS2_MYCGE	P47316	mycoplasma
C	33	6	3.3	290	1	YA57_METJA	O58457	methanococc
C	34	6	3.3	298	1	Y295_RICPR	O9zdn2	rickettsia
C	35	6	3.3	298	1	YDHF_ECOLI	P76187	escherichia
C	36	6	3.3	301	1	MCAT_HUMAN	O43772	homo sapien
C	37	6	3.3	301	1	MCAT_MOUSE	O9z226	mus musculu
C	38	6	3.3	307	1	KIFE_BORBU	O51575	borrelia bu
C	39	6	3.3	316	1	YQI5_CAEEL	O09278	caenorhabdl
C	40	6	3.3	318	1	DUS2_MOUSE	O05922	mus musculu
C	41	6	3.3	320	1	Y149_MYCPN	P75583	mycoplasma
C	42	6	3.3	326	1	NUIM_CHOCR	P48898	chondrus cr
C	43	6	3.3	326	1	RPOA_FUSNN	O8re44	fusobacteri
C	44	6	3.3	332	1	ETFA_CAEEL	O93615	caenorhabdl
C	45	6	3.3	332	1	G3P_BUCAI	P57384	buchnera ap

ALIGNMENTS

RESULT 1
RPOA_CAMJF
ID RPOA_CAMJF STANDARD; PRT; 337 AA.
AC Q9PM80;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase alpha chain (EC 2.7.7.6) (Transcriptase
DE alpha chain) (RNA polymerase alpha subunit).
GN RPOA OR C01595.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
-!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
-!- OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
-!- SUBSTRATES.
-!- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
-!- (RNA)(N).
-!- SUBUNIT: CONSISTS OF A SIGMA FACTOR AND THE RNAP CORE ENZYME WHICH
-!- IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, 1 BETA' CHAIN AND 1
-!- OMEGA CHAIN (BY SIMILARITY).
-!- DOMAIN: THE AMINO-TERMINAL PORTION IS INVOLVED IN THE ASSEMBLY OF
-!- CORE RNAP, WHEREAS THE C-TERMINAL IS INVOLVED IN INTERACTION WITH
-!- TRANSCRIPTIONAL REGULATORS (BY SIMILARITY).
-!- SIMILARITY: BELONGS TO THE RNA POLYMERASE ALPHA CHAIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AL139079; CAB73583.1; -.
DR HSSP; P00574; 1COO.
DR InterPro; IPR001700; RNA_pol_A_bac.
DR Pfam; PF01000; RNA_pol_A_bac; 1.
DR Pfam; PF03118; RNA_pol_A_CTD; 1.
DR ProDom; PD001179; RNA_pol_A_bac; 1.
KW Transferase; Transcription; DNA-directed RNA polymerase;
KW Complete proteome.
SQ SEQUENCE 337 AA; 37686 MW; 4BD2F59205DDE143 CRC64;

Alignment Scores:
Pred. No.: 10.8 Length: 337
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-22 (1-843) x RPOA_CAMJE (1-337)

QY 96 TTAATAAACCTGCAAAAAA 116
DB 290 LeuLysAsnLeuGlyLysLys 296

RESULT 2
HFEL_HAEIN
ID HFEL_HAEIN STANDARD; PRT; 435 AA.
AC P45994;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Minor fimbrial subunit hife precursor.
GN HIFE.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=Eagan / Serotype B;
RX MEDLINE=95012708; PubMed=7927773;
RA McCrea K.W., Watson W.J., Gilsdorf J.R., Mairs C.F.;
RT "Identification of hifD and hife in the pilus gene cluster of
RT Haemophilus influenzae type b strain Eagan.";
RL Infect. Immun. 62:4922-4928(1994).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=AM30 (770235) / Serotype B;
RX MEDLINE=95089703; PubMed=7997179;
RA van Ham M.S., van Alphen L., Mooi F.R., van Putten J.P.M.;
RT "The fimbrial gene cluster of Haemophilus influenzae type b.";
RL Mol. Microbiol. 13:673-684(1994).
CC -1- FUNCTION: MAY BE A MINOR STRUCTURAL PROTEIN REQUIRED FOR PILUS
CC BIOGENESIS. MAY BE THE ADHESIVE COMPONENT IN THE PILI.
CC -1- SIMILARITY: TO B.PERTUSSIS FIMD AND E.COLI FIMH.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U13254; AAA61542.1; -.
DR EMBL; Z33502; CAA83904.1; -.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.

KW Fimbria; Signal.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 435 MINOR FIMBRIAL SUBUNIT HIFE.
SQ SEQUENCE 435 AA; 48851 MW; 420C733661DE7D67 CRC64;

Alignment Scores:
Pred. No.: 10.3 Length: 435
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-22 (1-843) x HFEL_HAEIN (1-435)

QY 776 CTTAATAAACACCTTCCCC 796
DB 389 LeuLysLysGlnProSerPro 395

RESULT 3
VCAP_HSVSA
ID VCAP_HSVSA STANDARD; PRT; 1371 AA.
AC Q00999;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major capsid protein (MCP).
GN 25.
OS Herpesvirus saimiri (strain 11).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10383;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=92333688; PubMed=1321287;
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA Honess R.W.;
RT "Primary structure of the herpesvirus saimiri genome.";
RL J. Virol. 66:5047-5058(1992).
CC -1- FUNCTION: MAJOR PROTEIN OF THE ICOSAHEDRAL CAPSID.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES MAJOR CAPSID PROTEIN
CC FAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; X64346; CAA45648.1; -.
DR PIR; E36808; VCBEM7.
DR InterPro; IPR000912; Herpes_MCP.
DR Pfam; PF03122; Herpes_MCP; 1.
DR PRINTS; PR00235; HSVCAPSIDMCP.
KW Coat protein; Late protein.
SQ SEQUENCE 1371 AA; 154354 MW; 5F99FA9553A201C1 CRC64;

Alignment Scores:
Pred. No.: 8.15 Length: 1371
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-22 (1-843) x VCAP_HSVSA (1-1371)

QY 786 AACCTTCCGCCNCTTGGCC 806
DB 560 AsnLeuProProLeuAla 566


```
RESULT 4
Y031_TREPA
ID Y031_TREPA STANDARD; PRT; 92 AA.
AC 083074;
DT 15-DEC-1998 (Rel. 37, Last Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0031.
GN TP0031.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uitterback T.,
RA McDonald L., Artlisch P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001188; AAC65032.1; -.
DR TIGR: TP0031; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 65 POTENTIAL.
SQ SEQUENCE 92 AA; 10276 MW; 4EB638881276C4E8 CRC64;

Alignment Scores:
Pred. No.: 195 Length: 92
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-22 (1-843) x Y031_TREPA (1-92)

OY 115 TTTTTCNAGCTTTT 98
Db 30 PhepheProArGpPhephe 35

RESULT 5
ING_CHICK
ID ING_CHICK STANDARD; PRT; 164 AA.
AC P49708;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon gamma precursor (IFN-gamma).
GN IFNG.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96154849; PubMed=8590305;
```

```
RA Digby M.R., Lowenthal J.W.;
RT "Cloning and expression of the chicken interferon-gamma gene."
RL J. Interferon Cytokine Res. 15:939-945(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054661; PubMed=8898958;
RA Weining K.C., Schultz U., Muenster U., Kaspers B., Staeheli P.;
RT "Biological properties of recombinant chicken interferon-gamma."
RL Eur. J. Immunol. 26:2440-2447(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=N LINE;
RX MEDLINE=98172731; PubMed=9511740;
RA Kaiser P., Wain H.M., Rothwell L.;
RT "Structure of the chicken interferon-gamma gene, and comparison to
RT mammalian homologues."
RL Gene 207:25-32(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Spleen;
RA Song K.D., Kim D.K., Choi K.D., Lillehoj H.S., Han I.K., Han J.Y.;
RT "Molecular genetic analysis of chicken interferon gamma gene."
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PRODUCED BY LYMPHOCYTES ACTIVATED BY SPECIFIC ANTIGENS
CC OR MITOGENS. IFN-GAMMA, IN ADDITION TO HAVING ANTIVIRAL ACTIVITY,
CC HAS IMPORTANT IMMUNOREGULATORY FUNCTIONS. IT IS A POTENT ACTIVATOR
CC OF MACROPHAGES, IT HAS ANTIPROLIFERATIVE EFFECTS ON TRANSFORMED
CC CELLS AND IT CAN POTENTIATE THE ANTIVIRAL AND ANTITUMOR EFFECTS OF
CC THE TYPE I INTERFERONS (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE TYPE II (OR GAMMA) INTERFERON FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U27465; AAC59723.1; -.
DR EMBL: X99774; CAA68118.1; -.
DR EMBL: Y07922; CAA69227.1; -.
DR EMBL: U96875; AAG27918.1; -.
DR EMBL: U96872; AAG27918.1; JOINED.
DR EMBL: U96873; AAG27918.1; JOINED.
DR EMBL: U96874; AAG27918.1; JOINED.
DR HSSP: P07353; ID9C.
DR InterPro: IPR002069; IFN-gamma.
DR Pfam: PF00714; IFN-gamma; 1.
DR ProDom: PD002435; IFN-gamma; 1.
KW Cytokine; Antiviral; Growth regulation; Glycoprotein; Signal.
FT SIGNAL 19 POTENTIAL.
FT CHAIN 20 164 INTERFERON GAMMA.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 164 AA; 18953 MW; 18D9B0BCA415B600 CRC64;

Alignment Scores:
Pred. No.: 173 Length: 164
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-22 (1-843) x ING_CHICK (1-164)

OY 777 TTTAAAAACAACCTTCC 794
Db 103 LeuLysAsnAsnLeuPro 108
```

```
RESULT 6
ING_MELGA
ID   ING_MELGA          STANDARD;          PRT;          164 AA.
AC   057603;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon gamma precursor (IFN-gamma).
GN   IFNG.
OS   Meleagris gallopavo (Common turkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX   NCBI_TaxID=9103;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Kaiser P., Sonnemans D., Smith L.M.;
RT   "Avian interferon-gamma: cloning, sequencing and comparison of
RT   interferon-gamma genes from several different avian species.";
RL   Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
CC   -1- FUNCTION: PRODUCED BY LYMPHOCYTES ACTIVATED BY SPECIFIC ANTIGENS
CC   OR MITOGENS. IFN-GAMMA, IN ADDITION TO HAVING ANTIVIRAL ACTIVITY,
CC   HAS IMPORTANT IMMUNOREGULATORY FUNCTIONS. IT IS A POTENT ACTIVATOR
CC   OF MACROPHAGES, IT HAS ANTIPROLIFERATIVE EFFECTS ON TRANSFORMED
CC   CELLS AND IT CAN POTENTIATE THE ANTIVIRAL AND ANTITUMOR EFFECTS OF
CC   THE TYPE I INTERFERONS (BY SIMILARITY).
CC   -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -1- SIMILARITY: BELONGS TO THE TYPE II (OR GAMMA) INTERFERON FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AJ000725; CAA04260.1; -.
DR   HSSP; P07353; 1D9C.
DR   InterPro; IPR002069; IFN-gamma.
DR   Pfam; PF00714; IFN-gamma; 1.
DR   ProDom; PD002435; IFN-gamma; 1.
KW   Cytokine; Antiviral; Growth regulation; Glycoprotein; Signal.
FT   SIGNAL          1      19      POTENTIAL.
FT   CHAIN           20     164     INTERFERON GAMMA.
FT   CARBOHYD        42     42     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD        61     61     N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE      164 AA;  18943 MW;  3488F5A9A405B5C7 CRC64;

Alignment Scores:
Pred. No.:      173      Length:      164
Score:          6.00      Matches:      6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     3.33%      Indels:      0
DB:             1      Gaps:      0

US-09-835-992A-22 (1-843) x ING_MELGA (1-164)
QY   777 TTAATAAACACCTTCCC 794
      |||
Db   103 LeuLysAsnAsnLeuPro 108

RESULT 7
ING_PHACO
ID   ING_PHACO          STANDARD;          PRT;          164 AA.
AC   057608;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon gamma precursor (IFN-gamma).
GN   IFNG.
```

```
OS   Phasianus colchicus colchicus (Ring-necked pheasant).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Phasianus.
OX   NCBI_TaxID=9057;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kaiser P., Sonnemans D., Smith L.M.;
RT   "Avian interferon-gamma: cloning, sequencing and comparison of
RT   interferon-gamma genes from several different avian species.";
RL   Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
CC   -1- FUNCTION: PRODUCED BY LYMPHOCYTES ACTIVATED BY SPECIFIC ANTIGENS
CC   OR MITOGENS. IFN-GAMMA, IN ADDITION TO HAVING ANTIVIRAL ACTIVITY,
CC   HAS IMPORTANT IMMUNOREGULATORY FUNCTIONS. IT IS A POTENT ACTIVATOR
CC   OF MACROPHAGES, IT HAS ANTIPROLIFERATIVE EFFECTS ON TRANSFORMED
CC   CELLS AND IT CAN POTENTIATE THE ANTIVIRAL AND ANTITUMOR EFFECTS OF
CC   THE TYPE I INTERFERONS (BY SIMILARITY).
CC   -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -1- SIMILARITY: BELONGS TO THE TYPE II (OR GAMMA) INTERFERON FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AJ001289; CAA04649.1; -.
DR   HSSP; P07353; 1D9C.
DR   InterPro; IPR002069; IFN-gamma.
DR   Pfam; PF00714; IFN-gamma; 1.
DR   ProDom; PD002435; IFN-gamma; 1.
KW   Cytokine; Antiviral; Growth regulation; Glycoprotein; Signal.
FT   SIGNAL          1      19      POTENTIAL.
FT   CHAIN           20     164     INTERFERON GAMMA.
FT   CARBOHYD        42     42     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD        61     61     N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE      164 AA;  18910 MW;  4B943A819405B716 CRC64;

Alignment Scores:
Pred. No.:      173      Length:      164
Score:          6.00      Matches:      6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     3.33%      Indels:      0
DB:             1      Gaps:      0

US-09-835-992A-22 (1-843) x ING_PHACO (1-164)
QY   777 TTAATAAACACCTTCCC 794
      |||
Db   103 LeuLysAsnAsnLeuPro 108

RESULT 8
RL9_CHLPN
ID   RL9_CHLPN          STANDARD;          PRT;          169 AA.
AC   Q9Z6V3; Q9JQ51;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L9.
GN   RPLI OR RL9 OR CPN0953 OR CP0906.
OS   Chlamydia pneumoniae (Chlamydia pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Iamuel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
```

```
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AR39;
RX MEDLINE=20150255; Pubmed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weldman J., Knouft H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN-J138;
RX MEDLINE=20330349; Pubmed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- FUNCTION: BINDS TO THE 23S RNA (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L9P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001675; AAD19091.1; -.
CC EMBL; AE002249; AAF38691.1; -.
CC EMBL; AP002548; BAA99161.1; -.
CC PHCI-2DPAGE; Q926V3; -.
CC TIGR; CP0906; -.
CC InterPro; IPR000244; Ribosomal_L9.
CC Pfam; PF01281; Ribosomal_L9; 1.
CC TIGRFAMS; TIGR00158; L9; 1.
CC PROSITE; PS00651; RIBOSOMAL_L9; FALSE_NEG.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 169 AA; 18722 MW; AAA2F124F9EB4377 CRC64; .

Alignment Scores:
Pred. No.: 172 Length: 169
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x RL9_CHLPN (1-169)
QY 99 AAAAACCTNGGAAAAAA 116
Db 124 LysAsnLeuGlyLysLys 129

RESULT 9
APT_MASHI STANDARD; PRT; 180 AA.
AC Q64427;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
GN APRT.
OS Mastomys hildebrandtii.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Mastomys.
```

```
OX NCBI_TaxID=34847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97176443; Pubmed=9023989;
RA Fieldhouse D., Yazdani F., Golding G.B.;
RT "Substitution rate variation in closely related rodent species.";
RL Heredity 78:21-31(1997).
CC -1- FUNCTION: CATALYZES A SALVAGE REACTION RESULTING IN THE FORMATION
CC OF AMP, THAT IS ENERGICALLY LESS COSTLY THAN DE NOVO SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: AMP + diphosphate -> adenine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -1- PATHWAY: Purine salvage.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U28722; AAA68956.1; -.
CC InterPro; IPR000836; PRTtransferase.
CC InterPro; IPR002375; Pr/py_rp_transf.
CC Pfam; PF00156; Prbbsyltran; 1.
CC TIGRFAMS; TIGR01090; apt; 1.
CC PROSITE; PS00103; PUR_PYP_R_TRANSFER; 1.
KW Transferase; Glycosyltransferase; Purine salvage.
SQ SEQUENCE 180 AA; 19679 MW; 07E76B24159FF099 CRC64; .

Alignment Scores:
Pred. No.: 170 Length: 180
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x APT_MASHI (1-180)
QY 771 GGACNATCCCGTTTTC 754
Db 169 GlyProIleProPhePhe 174

RESULT 10
APT_MOUSE STANDARD; PRT; 180 AA.
AC P08030;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
GN APRT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190571; Pubmed=3921964;
RA Dush M.K., Sikela J.M., Khan S.A., Tischfield J.A., Stambrook P.J.;
RT "Nucleotide sequence and organization of the mouse adenine
RT phosphoribosyltransferase gene: presence of a coding region common to
RT animal and bacterial phosphoribosyltransferases that has a variable
RT intron/exon arrangement.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2731-2735(1985).
CC -1- FUNCTION: CATALYZES A SALVAGE REACTION RESULTING IN THE FORMATION
CC OF AMP, THAT IS ENERGICALLY LESS COSTLY THAN DE NOVO SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: AMP + diphosphate -> adenine + 5-phospho-
```

```
CC      alpha-D-ribose 1-diphosphate.
CC      -1- PATHWAY: Purine salvage.
CC      -1- SUBUNIT: HOMODIMER.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC      PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M11310; AAA37255.1; -.
DR      PIR; A22670; RTMSA.
DR      MGD; MGI:88061; APT.
DR      InterPro; IPR000836; PRtransferase.
DR      InterPro; IPR002375; PR/py_rp_transf.
DR      Pfam; PF00156; Pribosyltran; 1.
DR      TIGRFAMS; TIGR01090; apt; 1.
DR      PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
KW      Transferase; Glycosyltransferase; Purine salvage.
SQ      SEQUENCE 180 AA; 19736 MW; 8E4F0CBA1F173C64 CRC64; .

Alignment Scores:
Pred. No.:      170      Length:      180
Score:          6.00      Matches:      6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    3.28%      Indels:      0
DB:             1      Gaps:          0

US-09-835-992A-22 (1-843) x APT_MOUSE (1-180)

QY      771 GGACCNATCCGTTTTC 754
      11111111111111111111
Db      169 GlyProileProphphe 174

RESULT 11
APT_MUSPA
ID      APT_MUSPA      STANDARD;      PRT;      180 AA.
AC      P47956;
DT      01-FEB-1996 (Rel. 33, Created)
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
GN      APRT.
OS      Mus pahari (Shrew mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10093;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97176443; PubMed=9023989;
RA      Fieldhouse D., Yazdani F., Golding G.B.;
RT      "Substitution rate variation in closely related rodent species.";
RL      Heredity 78:21-31(1997).
CC      -1- FUNCTION: CATALYZES A SALVAGE REACTION RESULTING IN THE FORMATION
CC      OF AMP, THAT IS ENERGETICALLY LESS COSTLY THAN DE NOVO SYNTHESIS.
CC      -1- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
CC      alpha-D-ribose 1-diphosphate.
CC      -1- PATHWAY: Purine salvage.
CC      -1- SUBUNIT: HOMODIMER.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC      PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
```

```
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U28721; AAA68957.1; -.
DR      InterPro; IPR000836; PRtransferase.
DR      InterPro; IPR002375; PR/py_rp_transf.
DR      Pfam; PF00156; Pribosyltran; 1.
DR      TIGRFAMS; TIGR01090; apt; 1.
DR      PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
KW      Transferase; Glycosyltransferase; Purine salvage.
SQ      SEQUENCE 180 AA; 19712 MW; A8F43EB56F4D2CBF CRC64; .

Alignment Scores:
Pred. No.:      170      Length:      180
Score:          6.00      Matches:      6
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    3.28%      Indels:      0
DB:             1      Gaps:          0

US-09-835-992A-22 (1-843) x APT_MUSPA (1-180)

QY      771 GGACCNATCCGTTTTC 754
      11111111111111111111
Db      169 GlyProileProphphe 174

RESULT 12
APT_MUSSI
ID      APT_MUSSI      STANDARD;      PRT;      180 AA.
AC      P47957;
DT      01-FEB-1996 (Rel. 33, Created)
DT      01-FEB-1996 (Rel. 33, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
GN      APRT.
OS      Mus spicilegus (Steppe mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10103;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97176443; PubMed=9023989;
RA      Fieldhouse D., Yazdani F., Golding G.B.;
RT      "Substitution rate variation in closely related rodent species.";
RL      Heredity 78:21-31(1997).
CC      -1- FUNCTION: CATALYZES A SALVAGE REACTION RESULTING IN THE FORMATION
CC      OF AMP, THAT IS ENERGETICALLY LESS COSTLY THAN DE NOVO SYNTHESIS.
CC      -1- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
CC      alpha-D-ribose 1-diphosphate.
CC      -1- PATHWAY: Purine salvage.
CC      -1- SUBUNIT: HOMODIMER.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC      PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
```


Alignment Scores:

Pred. No.:	170	Length:	180
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.28%	Indels:	0
DB:	1	Gaps:	0

US-09-835-992a-22 (1-843) x APT_MUSSI (1-180)

OY 771 GGACCNATCCGTTTTC 754
 DB 169 GlyProileProPhe 174

RESULT 13

RRF_DEIRA STANDARD; PRT; 183 AA.

AC 09R082;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ribosome recycling factor (Ribosome releasing factor) (RRF).
 GN FRR OR DR1510.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID-1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-R1;
 RX MEDLINE-20036896; PubMed-10567266;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RA "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1.";
 RL Science 286:1571-1577(1999).

CC -1- FUNCTION: Responsible for the release of ribosomes from messenger
 CC RNA at the termination of protein biosynthesis. May increase the
 CC efficiency of translation by recycling ribosomes from one round of
 CC translation to another (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE RRF FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE001995; AAF11077.1; -
 DR HSSP; Q9X1B9; IDD5.
 DR TIGR; DR1510; -
 DR InterPro; IPR002661; RRF.
 DR Pfam; PF01765; RRF; 1.
 DR TIGRFAMs; TIGR00496; frr; 1.
 KW Protein biosynthesis; Complete proteome.
 SQ SEQUENCE 183 AA; 20620 MW; 40B1546728CE5B57 CRC64;

Alignment Scores:
 Pred. No.: 170
 Score: 6.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 3.33%
 DB: 1

Length: 183
 Matches: 6
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-835-992a-22 (1-843) x RRF_DEIRA (1-183)

OY 823 CTAAATCCGAACATATA 840
 DB 91 LeuAsnProAsnAsnLys 96

RESULT 14

YPDC_BACSU STANDARD; PRT; 218 AA.

AC P50738;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein ypdC.
 GN YPDC.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID-1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168 / Marburg;
 RX MEDLINE-96349105; PubMed-8760912;
 RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
 RA Serror P.;
 RA "Sequence analysis of the Bacillus subtilis chromosome region between
 RT the sera and kdg loci cloned in a yeast artificial chromosome.";
 RL Microbiology 142:2005-2016(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RX MEDLINE-98044033; PubMed-9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borris R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Fougere D.,
 RA Fritze C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Jorais B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield E.,
 RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weltzenegger T.,
 RA Winters P., Wilpat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; LA7648; AAC83956.1; -

DR EMBL; 299115; CAB14210.1; -
DR EMBL; 299116; CAB14226.1; -
DR Subtilist; BG11438; YPdc.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 127 147 POTENTIAL.
FT TRANSMEM 180 200 POTENTIAL.
SQ SEQUENCE 218 AA; 24719 MW; 88556D50863E14BC CRC64;

Alignment Scores:
Pred. No.: 164 Length: 218
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-22 (1-843) x YPDC_BACSU (1-218)

QY 676 TTTTACCTTGCTAAGGCN 693
Db 147 PheTyrLeuGlyLysAla 152

RESULT 15
YRAL_YEAST
ID YRAL_YEAST STANDARD; PRT; 226 AA.
AC Q12159;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE RNA annealing protein YRAL.
GN YRAL OR YDR381W OR D9481.2 OR D9509.1.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 39-53 AND 57-70.
RC STRAIN=Bj5626;
RX MEDLINE=97293180; PubMed=9149233;
RA Portman D.S., O'Connor J.P., Dreyfuss G.;
RT "YRAL, an essential Saccharomyces cerevisiae gene, encodes a novel
nuclear protein with RNA annealing activity.";
RL RNA 3:527-537(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,
RA Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,
RA Taich A., Trevasakis E., Vignati D., Wilcox L., Woldman P., Vaudin M.,
RA Wilson R., Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
RA Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA Hunnicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA Winant A., Yelton M., Botstein D., Davis R.W.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RNA-BINDING RNA ANNEALING PROTEIN. MAY HAVE A ROLE IN
PRE-MRNA METABOLISM.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -!- SIMILARITY: TO S.POMBE MLO3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U72633; AAC09951.1; -
DR EMBL; U28373; AAB64817.1; -
DR EMBL; U32274; AAB64823.1; -
DR SGD; S0002789; YRAL.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW Nuclear protein; RNA-binding.
FT DOMAIN 78 158 RNA-BINDING (RRM).
FT DOMAIN 201 210 ARG/LYS-RICH (BASIC).
SQ SEQUENCE 226 AA; 24955 MW; 480B5B6DC0D14BE9 CRC64;

Alignment Scores:
Pred. No.: 162 Length: 226
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-22 (1-843) x YRAL_YEAST (1-226)

QY 619 TTTACGGGTCCCAATT 636
Db 139 PheAsnGlySerProile 144

Search completed: January 14, 2003, 17:29:21
Job time : 17.0173 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:20:59 ; Search time 59.3589 Seconds
(without alignments)
5852.461 Million cell updates/sec

Title: US-09-835-992A-22
Perfect score: 180
Sequence: 1 ggcccaaaaantattctna.....taaatccgaacaataaaag 843

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+_n2p.model -DEV-xlp
-Q-/cgn2_1/USPTO_spool/US09835992/runat_14012003_161526_4027/app_query.fasta_1.3932
-DB-SPTREMBL_21 -QFMT-fastan -SUFFIX-olig.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.cdt
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER-US09835992_@cgn_1_1_346_@runat_14012003_161526_4027 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	8	4.4	288	5 Q95QI1	Q95q11 caenorhabdi

C	2	7	3.8	86	10	Q41572	Q41572 triticum ae
C	3	7	3.9	172	16	O53177	O53177 mycobacteri
C	4	7	3.9	188	16	O8VJ14	O8VJ14 mycobacteri
C	5	7	3.9	215	5	O9BL14	O9BL14 papilio xut
C	6	7	3.9	223	10	Q98S75	Q98S75 guillardia
C	7	7	3.9	263	10	Q9FR36	Q9FR36 glycine max
C	8	7	3.9	328	2	O9F5H2	O9F5H2 agrobacteri
C	9	7	3.9	387	9	O03951	O03951 bacterioph
C	10	7	3.9	426	6	O9GKN7	O9GKN7 bos taurus
C	11	7	3.9	435	2	O99Q04	O99Q04 haemophilus
C	12	7	3.9	435	2	O9AIR2	O9AIR2 haemophilus
C	13	7	3.9	498	10	Q9M2D7	Q9M2D7 arabidopsis
C	14	7	3.9	1240	3	Q9P6U5	Q9P6U5 neurospora
C	15	7	3.9	1371	12	O9YTP1	O9YTP1 atelina her
C	16	7	3.9	1379	12	Q993J3	Q993J3 callitrichl
C	17	6	3.3	25	13	P82741	P82741 rana catesb
C	18	6	3.3	25	13	P82875	P82875 rana clamit
C	19	6	3.3	51	11	P97645	P97645 rattus norv
C	20	6	3.3	63	2	Q9AM38	Q9AM38 yersinia ps
C	21	6	3.3	68	16	O9JV59	O9JV59 neisseria m
C	22	6	3.3	75	2	O54661	O54661 streptococc
C	23	6	3.3	75	12	O8QL99	O8QL99 mamestra co
C	24	6	3.3	76	10	O22492	O22492 oenanthhe ja
C	25	6	3.3	87	17	O8TTK7	O8TTK7 methanosarc
C	26	6	3.3	90	5	O9XXC2	O9XXC2 caenorhabdi
C	27	6	3.3	90	12	O64857	O64857 human adeno
C	28	6	3.3	90	16	O9HV94	O9HV94 pseudomonas
C	29	6	3.3	91	2	O8VL47	O8VL47 escherichia
C	30	6	3.3	92	5	O9V357	O9V357 drosophila
C	31	6	3.3	94	2	O54670	O54670 streptococc
C	32	6	3.3	95	9	O9AZU4	O9AZU4 bacterioph
C	33	6	3.3	95	16	O9CFQ6	O9CFQ6 lactococcus
C	34	6	3.3	110	10	O9LZD1	O9LZD1 arabidopsis
C	35	6	3.3	116	8	O8W9S0	O8W9S0 mesostigma
C	36	6	3.3	117	10	O40553	O40553 nicotiana t
C	37	6	3.3	119	2	O46301	O46301 clostridium
C	38	6	3.3	125	10	O9AW17	O9AW17 guillardia
C	39	6	3.3	126	4	O12444	O12444 saccharomyc
C	40	6	3.3	127	4	O9HAF8	O9HAF8 homo sapien
C	41	6	3.3	127	8	O98RP5	O98RP5 guillardia
C	42	6	3.3	129	15	O36880	O36880 human immun
C	43	6	3.3	131	13	O90276	O90276 boulangerei
C	44	6	3.3	135	4	O9H5Z5	O9H5Z5 homo sapien
C	45	6	3.3	136	13	O91254	O91254 paracheitrod

ALIGNMENTS

RESULT 1	ID	Q95QI1	PRELIMINARY;	PRT;	288 AA.
AC	Q95QI1;				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	F30F8.9 protein.				
GN	F30F8.9.				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;				
OC	Rhabditidae; Peloderinae; Caenorhabditis.				
OX	NCBI_TaxID=6239;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Wilkinson J.;				
RL	Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99069613; PubMed=9851916;				
RA	none;				
RT	"Genome sequence of the nematode C.elegans: A platform for				
RT	investigating biology."				
RL	Science 282:2012-2018(1998).				
DR	EMBL; Z81073; CAC42296.1; -.				

SQ SEQUENCE 288 AA; 33108 MW; 06441683774F7B14 CRC64;

Alignment Scores:

Pred. No.:	1.86	Length:	288
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.37%	Indels:	0
DB:	5	Gaps:	0

US-09-835-992A-22 (1-843) x Q95Q11 (1-288)

OY 127 TTTCNATATTTTTCNAGG 104
|||||
Db 131 PheAlaIleIlePheProArg 138

RESULT 2
Q41572
ID Q41572 PRELIMINARY; PRT; 86 AA.
AC Q41572;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE (Begra) ORF1 and ORF2 genes.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BEGRA;
RA Kroczyńska B., Buchowicz J.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; X76181; CAA53774.1; -.
SQ SEQUENCE 86 AA; 9783 MW; 82D92DDDBD138242 CRC64;

Alignment Scores:

Pred. No.:	37.6	Length:	86
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.83%	Indels:	0
DB:	10	Gaps:	0

US-09-835-992A-22 (1-843) x Q41572 (1-86)

OY 799 GGNGGGGAGCTGTTT 779
|||||
Db 37 GlyGlyGlyArgLeuPhe 43

RESULT 3
O53177
ID O53177 PRELIMINARY; PRT; 172 AA.
AC O53177;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein RV2450c.
GN RV2450C OR MTV008.06C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham D., Gents M., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL; AL021246; CAA16027.1; -.
DR TubercuList; RV2450c; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 172 AA; 17456 MW; 0D67E72240EF0F0E CRC64;

Alignment Scores:

Pred. No.:	33.2	Length:	172
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.89%	Indels:	0
DB:	16	Gaps:	0

US-09-835-992A-22 (1-843) x O53177 (1-172)

OY 786 AACCTTCCCCNCCTTGCC 806
|||||
Db 73 AsnLeuProProLeuAla 79

RESULT 4
Q8VJ14
ID Q8VJ14 PRELIMINARY; PRT; 188 AA.
AC Q8VJ14;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein MT2526.
GN MT2526.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weldman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE007090; AAK46825.1; -.
DR TIGR; MT2526; -.
KW Hypothetical protein.
SQ SEQUENCE 188 AA; 19223 MW; 3AF2AD19B97736E2 CRC64;

Alignment Scores:

Pred. No.:	32.6	Length:	188
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.89%	Indels:	0
DB:	16	Gaps:	0

US-09-835-992A-22 (1-843) x Q8VJ14 (1-188)

OY 786 AACCTTCCCCNCCTTGCC 806
|||||
Db 89 AsnLeuProProLeuAla 95

RESULT 5
Q9BL14
ID Q9BL14 PRELIMINARY; PRT; 215 AA.
AC Q9BL14;
DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DE Fibroin P25.
GN FIB-P25.
OS Papilio xuthus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Papilionoidea; Papilionidae; Papilioninae; Papilio.
OX NCBI_TaxID=66420;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21167527; PubMed=11267905;
RA Tanaka K., Mizuno S.;
RT "Homologues of fibroin L-chain and P25 of Bombyx mori are present in
RT Dendrolimus spectabilis and Papilio xuthus but not detectable in
RT Anthraea yamamai.";
RL Insect Biochem. Mol. Biol. 31:665-677(2001).
DR EMBL; AB001825; BAB39504.1; -.
SQ SEQUENCE 215 AA; 24346 MW; 21BD9C93FDF83389 CRC64;

Alignment Scores:
Pred. No.: 31.8 Length: 215
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 5 Gaps: 0

US-09-835-992A-22 (1-843) x Q9BLI4 (1-215)

OY 820 CGCTTAATCCGAACAATAA 840
Db 165 ArgLeuAsnProAsnAsnLys 171

RESULT 6

O98S75 PRELIMINARY; PRT; 223 AA.

AC O98S75;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)
DE Hypothetical 26.9 kDa protein orf223 from chromosome 3.
GN ORF223.
OS Guillardia theta (Cryptomonas ph1).
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX NCBI_TaxID=55529;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21223349; PubMed=11323671;
RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
RA Wu X., Reith M., Cavalier-Smith T., Mäler U.G.;
RT "The highly reduced genome of an enslaved algal nucleus."
RL Nature 410:1091-1096(2001).
DR EMBL; AF083031; AAK39707.1; -.
KW Hypothetical protein.
SQ SEQUENCE 223 AA; 26902 MW; A0396B214866F092 CRC64;

Alignment Scores:
Pred. No.: 31.6 Length: 223
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x Q98S75 (1-223)

OY 99 AAAAACCTNGCAAAAAATA 119
Db 87 LysAsnLeuGlyLysLysile 93

RESULT 7
O9FR36

ID O9FR36 PRELIMINARY; PRT; 263 AA.
AC O9FR36;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Putative steroid reductase.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. WILLIAMS 82; TISSUE=LEAF;
RX MEDLINE=99169286; PubMed=10069834;
RA Reverdatto S., Bellinson V., Nielsen N.C.;
RT "A multisubunit acetyl coenzyme A carboxylase from soybean."
RL Plant Physiol. 119:961-978(1999).
DR EMBL; AF203341; AAG35638.1; -.
DR InterPro; IPR001104; Strd5A_dhc.
DR Pfam; PF02544; Steroid_dh; 1.
SQ SEQUENCE 263 AA; 30263 MW; 22DC8E76DD501135 CRC64;

Alignment Scores:
Pred. No.: 30.7 Length: 263
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x Q9FR36 (1-263)

OY 786 AACCTTCCCCCNCCTTGCC 806

Db 48 AsnLeuProProProlLeuAla 54

RESULT 8

O9F5H2 PRELIMINARY; PRT; 328 AA.

AC O9F5H2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE Riorf20 protein.
GN RIORF20.
OS Agrobacterium rhizogenes.
OG Plasmid pRI1724.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF03-01724;
RA Moriguchi K., Maeda Y., Satou M., Satuti N., Kataoka M., Tanaka N.,
RA Yoshida K.;
RT "The complete nucleotide sequence of a Ri (root inducing) plasmid
RT indicates its chimerical structure between T1 and Sym plasmids."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF03-01724;
RA Moriguchi K., Maeda Y., Satou M., Kataoka M., Tanaka N., Yoshida K.;
RT "Analysis of unique variable region of a plant root inducing plasmid,
RT pRI1724, by the construction of its physical map and library."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF03-01724;
RA Moriguchi K., Nishida T., Maeda Y., Tanaka N., Yoshida K.;
RT "Genome structure of Ri plasmid (1):Construction of linking library
RT and physical map of pRI1724 in Japanese Agrobacterium.";
RL Nucleic Acids Symp. Ser. 39:189-190(1998).

```

RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF03-01724;
RX MEDLINE=20241294; PubMed=10780382;
RA Maeda Y., Moriguchi K., Kataoka M., Satou M., Satuti N., Tanaka N.,
RA Yoshida K.;
RT "Genome structure of R1 plasmid (1): Sequencing analysis of T-DNA and
RT its flanking regions of pR11724 in Japanese Agrobacterium
RT rhizogenes.";
RL Nucleic Acids Symp. Ser. 42:67-68(1999).
DR EMBL; AP002086; BAB16139.1; -.
KW Plasmid.
SQ SEQUENCE 328 AA; 36666 MW; 0632204AA307C2A5 CRC64;

Alignment Scores:
Pred. No.: 29.5 Length: 328
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x Q9F5H2 (1-328)

QY 673 CCATTTCCTTGTAAGGCN 693
Db 304 ProPhetYrLeuGlyLysAla 310

RESULT 9
O03951 PRELIMINARY; PRT; 387 AA.
AC 003951;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE GP28 (Partition protein A).
GN GENE 28 OR PARA.
OS Bacteriophage N15.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=40631;
RN [1]
RP SEQUENCE FROM N.A.
RA Hendrix R.W., Ravin V.K., Casjens S.R., Ford M.E., Ravin N.V.,
RA Smirnov I.K.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-207 FROM N.A.
RA Svarchevsky A.N., Rybchin V.N.;
RL Mol. Gen. Microbiol. Virusol. 2:16-22(1984).
RN [3]
RP SEQUENCE OF 1-207 FROM N.A.
RA Vostrov A.A., Iu Malinin A., Rybchin V.N., Svarchevsky A.N.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF064539; AAC19064.1; -.
DR EMBL; U91583; AAB58384.1; -.
DR InterPro; IPR000707; ATPase_Para.
DR Pfam; PF00991; Para; 1.
SQ SEQUENCE 387 AA; 43607 MW; 40FD9058078F992F CRC64;

Alignment Scores:
Pred. No.: 28.6 Length: 387
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 9 Gaps: 0

US-09-835-992A-22 (1-843) x O03951 (1-387)

QY 670 TTACCAATTTCCTTGTAAG 690
Db 169 LeuProPhetYrLeuGlyLys 175
```

```

RESULT 10
Q9GKN7 PRELIMINARY; PRT; 426 AA.
AC Q9GKN7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Rh50-like protein.
GN RH50.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Cherif-Zahar B., Raynal V., Cartron J.P., Matassi G.;
RT "Evolution of Rh50 genes in Metazoa.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF164575; AAG38955.1; -.
DR InterPro; IPR001905; Ammonium_transp.
DR InterPro; IPR002229; RhesusRHD.
DR InterPro; IPR000169; SHprot_acsite.
DR Pfam; PF00909; Ammonium_transp; 1.
DR PRINTS; PR00342; RHEUSRHD.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; UNKNOWN_1.
FT VARIANT 201 201 S -> P.
SQ SEQUENCE 426 AA; 46719 MW; C4899A8406E09E51 CRC64;

Alignment Scores:
Pred. No.: 28.1 Length: 426
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 6 Gaps: 0

US-09-835-992A-22 (1-843) x Q9GKN7 (1-426)

QY 409 TTTAAATTTTAACCCCTTA 429
Db 313 PhelysPheluThrProleu 319

RESULT 11
Q99Q04 PRELIMINARY; PRT; 435 AA.
AC Q99Q04;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE H1fE.
GN H1fE.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1595, 6351, AND 785;
RA Clemans D.L., Patel M.J., Marrs C.F., Bauer R.J., Gilsdorf J.R.;
RT "Analysis of the pilus adhesins from Haemophilus influenzae biotype IV
RT strains.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF245361; AAK28377.1; -.
DR EMBL; AF245359; AAK28375.1; -.
DR EMBL; AF245360; AAK28376.1; -.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
SQ SEQUENCE 435 AA; 48807 MW; E2E157E3E018677A CRC64;
```

```

Alignment Scores:
Pred. No.: 28 Length: 435
```


Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x Q99Q04 (1-435)

OY 776 CTTAAAAACAACCTTCCCC 796
| | | | | | | | | | | | | | | | | |
DB 389 LeuLySlyGlnProSerPro 395

RESULT 12

O9AIR2 PRELIMINARY; PRT; 435 AA.
AC O9AIR2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE H1FE.
CN H1FE.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1216;
RA Clemons D.L., Patel M.J., Marrs C.F., Bauer R.J., Gilsdorf J.R.;
RT "Analysis of the plus adhesins from Haemophilus influenzae biotype IV
strains.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF245363; AAK28379.1; -
DR InterPro: IPR000259; Fimbrin.
DR Pfam; PF00419; Fimbrin; 1.
SQ SEQUENCE 435 AA; 48767 MW; 477D8E92A749CE4D CRC64;

Alignment Scores:

Pred. No.: 28 Length: 435
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x Q9AIR2 (1-435)

OY 776 CTTAAAAACAACCTTCCCC 796
| | | | | | | | | | | | | | | | | |
DB 389 LeuLySlyGlnProSerPro 395

RESULT 13

O9M2D7 PRELIMINARY; PRT; 498 AA.
AC O9M2D7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE Hypothetical 58.5 kDa protein.
GN T20K12.170.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,
RA Mayer K.F.X., Quettier F., Salanoubat M.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137898; CAB71057.1; -
KW Hypothetical protein.
SQ SEQUENCE 498 AA; 58523 MW; B2E8606256BB8E4A CRC64;

Alignment Scores:

Pred. No.: 27.3 Length: 498
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-22 (1-843) x Q9M2D7 (1-498)

OY 817 TCCCGTCTAAATCCGAACAAT 837
| | | | | | | | | | | | | | | | | |
DB 84 SerArgLeuAsnProAsnAsn 90

RESULT 14

O9P6U5 PRELIMINARY; PRT; 1240 AA.
AC O9P6U5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Related to protease ULP2 protein.
GN 15E6.80.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Algn V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL353822; CAB88639.1; -
DR InterPro: IPR003653; SUMO_protease.
DR Pfam; PF02902; Peptidase_C48; 1.
KW Protease.
SQ SEQUENCE 1240 AA; 138114 MW; 716E38F4DF0D177A CRC64;

Alignment Scores:

Pred. No.: 23.1 Length: 1240
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 3 Gaps: 0

US-09-835-992A-22 (1-843) x Q9P6U5 (1-1240)

OY 767 GGTCCCCCCTTAAAAACAA 787
| | | | | | | | | | | | | | | | | |
DB 29 GlyProProLeuLySlySln 35

RESULT 15

O9YTP1 PRELIMINARY; PRT; 1371 AA.
AC O9YTP1;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Major capsid protein.
OS Ateeline herpesvirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=85618;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-73;
RX MEDLINE=20091363; PubMed=10623770;
RA Albrecht J.C.;
RT "Primary structure of the Herpesvirus Ateles genome.";
RL J. Virol. 74:1033-1037(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-73;
RA Albrecht J.-C., Fleckenstein B.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083424; AAC95549.1; -
DR InterPro; IPR000912; Herpes_MCP.
DR Pfam; PF03122; Herpes_MCP; 1.
DR PRINTS; PR00235; HSVCAPSIDMCP.
SQ SEQUENCE 1371 AA; 153800 MW; EE9FC518FDF066A0 CRC64;

Alignment Scores:
Pred. No.: 22.7 Length: 1371
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 12 Gaps: 0

US-09-835-992A-22 (1-843) x Q9YTP1 (1-1371)

OY 786 AACCTTCCCCCNCCTTGGCC 806
|||||
DB 560 AsnLeuProProLeuAla 566

Search completed: January 14, 2003, 17:37:09
Job time : 65.3589 secs